

83779

STIC-Biot ch/ChemLib

Fr m: Chan, Christina
S nt: Tuesday, January 07, 2003 1:14 PM
T : Duffy, Patricia; STIC-Biotech/ChemLib
Subject: RE: Please Rush -- Amendment Due this Biweek

Imp rtance: High

Pl ase rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: **Duffy, Patricia**
Sent: Tuesday, January 07, 2003 1:12 PM
To: Chan, Christina
Subject: Please Rush -- Amendment Due this Biweek
Importance: High

In re: 09/900,766

Please search SEQ ID NOs:1 and 2.
Please perform an interference search.
Please print out top 50 hits in each category.
Thanks mucho.

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Patricia A. Duffy
CM1-8D05
AU 1645
703-305-7555

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 1/8/03
Date Completed: 1/18/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: ✓ _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: PK _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:52:01 ; Search time 49.0077 Seconds
(without alignments)
1827.149 Million cell updates/sec

Title: US-09-900-766-1

Perfect score: 3522

Sequence: 1 EVQLQQSGDLVKPGASYKI.....EATHKTSTPIVKSFNRNES 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1644	46.7	573	22 AAG64459	Mouse 6D9 catalyti
2	1113	31.6	245	18 AAW33375	Staphylococcus ent
3	1107	31.4	257	22 AAU14103	Peptide sequence f
4	1096.5	31.1	720	22 AAG65572	Amino acid sequenc
5	1093	31.0	230	22 AAB67339	Staphylococcus aur
6	1091	31.0	245	18 AAW35374	Staphylococcus ent
7	1069	30.4	592	22 AAW38338	Amino acid sequenc
8	1069	30.4	595	22 AAW86003	Anti-5T4 single ch
9	1059.5	30.1	464	16 AAR76088	MAB 55.1 heavy cha
10	1055.5	30.0	445	16 AAR76085	MAB 55.1 heavy cha

11	1048	29.8	230	14 AAR45012	Staphylococcal ent
12	1044	29.6	230	23 ABB76235	Staphylococcus aur
13	1035	29.4	230	12 AAR13204	Staphylococcal ent
14	1028	29.2	597	23 AAE18377	Human N-terminal D
15	1025.5	29.1	464	19 AAW83041	Anti-Fas MAB HFE7A
16	1025.5	29.1	464	21 AAB14747	Mouse anti-Fas ant
17	1025.5	29.1	464	21 AAW90897	Murine anti-Fas an
18	1025.5	29.1	464	23 ABB74866	Humanised anti-Fas
19	1025.5	29.1	464	23 ABB74912	Humanised anti-Fas
20	1024.5	29.1	613	23 AAE18380	Human N-terminal D
21	1023.5	29.1	456	23 AAE18370	Human penton base
22	1023.5	29.1	493	23 AAE18379	Human N-terminal D
23	1023.5	29.1	510	23 AAE18378	Human N-terminal D
24	1022.5	29.0	438	23 AAE18372	Human penton base
25	1018	28.9	206	20 AAY39452	Antibody ABX-CBL 1
26	1018	28.9	465	16 AAR66758	Anti-tobacco mosai
27	980.5	27.8	711	20 AAW85692	MoTabII fusion pro
28	977.5	27.8	626	21 AAY55081	Single chain Fv pr
29	972.5	27.6	243	20 AAY42294	Anti-5T4 secreted
30	972.5	27.6	243	20 AAY27407	5T4 scFv antibody
31	972.5	27.6	243	20 AAW86002	Murine anti-5T4 an
32	972.5	27.6	243	22 AAB83835	Amino acid sequenc
33	970	27.5	214	20 AAY44176	MAB Fab1365 heavy
34	967.5	27.5	488	20 AAW86004	Human B7-1.5T4.1 p
35	967.5	27.5	488	22 AAB83836	Amino acid sequenc
36	966	27.4	468	12 AAR13061	Monoclonal antibody
37	961.5	27.3	225	14 AAR40385	Monoclonal antibody
38	960	27.3	212	17 AAW15932	Antibody 7G12 heav
39	960	27.3	233	12 AAR13203	Staphylococcal ent
40	953.5	27.1	223	16 AAR75456	Mouse antibody FB3
41	953	27.1	220	15 AAR53802	FAB light chain fo
42	951.5	27.0	469	14 AAR40384	Monoclonal antibody
43	948	26.9	214	21 AAU78253	Mouse agglutinatio
44	948	26.9	257	22 AAU14104	Peptide sequence f
45	944.5	26.8	239	16 AAR76087	MAB 55.1 light cha

ALIGNMENTS

```

RESULT 1
AAG64459
ID AAG64459 standard; Protein; 573 AA.
XX
AC AAG644459;
XX
DT 21-SEP-2001 (first entry)
XX
DE Mouse 6D9 catalytic antibody.
XX
KW Mouse; 6D9; catalytic antibody; esterase activity.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 28
FT FT /note= "Encoded by GCCTGA"
FT FT Misc-difference 35
FT FT /label= unknown
FT FT /note= "Encoded by TNT"
FT FT Misc-difference 46
FT FT /note= "Encoded by GAGTAA"
FT FT Misc-difference 329
FT FT /note= "Encoded by GAGTAA"
FT FT Misc-difference 572
FT FT /note= "Encoded by TGTAA"
XX
JP2001128681-A.
XX
15-MAY-2001.
XX
04-NOV-1999; 99JP-0314259.
XX
XX

```



```
XX SQ Sequence 245 AA;
Query Match 31.6%; Score 1113; DB 18; Length 245;
Best Local Similarity 86.9%; Pred. No. 2.5e-59;
Matches 213; Conservative 8; Mismatches 12; Indels 12; Gaps 1;

Qy 226 SEKSEINEKDLRKSELOQTALGNLKOIYYYNKAITSSSEKSAQDQFLNTLLFKGFFTG 285
Db 1 SEKSEINEKDLRKSELOQTALGNLKOIYYYNKAITENKESDDQFLNTLLFKGFFTG 60

Qy 286 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 345
Db 61 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120

Qy 346 BEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 405
Db 121 BEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180

Qy 406 RGLIVFHSSEGSTVSVDLFDAGGQVPTDQLLRIYRDN-----TISSTSLSL 453
Db 181 RGLIVFHSSEGSTVSVDLFDAGGQVPTDQLLRIYRDNKTINSENHLIDLYTINSENHLIDL 240

Qy 454 YLYTT 458
Db 241 YLYTT 245

RESULT 3
ID AAU14103 standard; peptide; 257 AA.
XX AC AAU14103;
XX DT 21-NOV-2001 (first entry)
XX DE Peptide sequence from Staphylococcus aureus enterotoxin type E.
XX KW Anti-retroviral; DP178-like; DP107-like; enterotoxin type E;
XX KW antifusogenic; antiviral; HIV transmission.
XX OS Staphylococcus aureus.
XX PN W0200151673-A2.
XX PD 19-JUL-2001.
XX PF 05-JUL-2000; 2000WO-US35727.
XX PR 09-JUL-1999; 99US-0350841.
XX PA (TRIM-) TRIMERIS INC.
XX PI Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
XX WPI; 2001-442157/47.
XX DR
XX PT Identifying a compound that inhibits the formation of or disrupts a
XX PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral
XX PT or intracellular modulatory activity, by detecting the formation of a
XX PT DP107/DP178 complex -
XX PS Disclosure; Fig 41; 259pp; English.
XX CC The present invention relates to peptides which exhibit anti-retroviral
XX CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
XX CC DP178-like and DP107-like peptides. The DP178 peptide corresponds
XX CC to amino acids 639-673 of the transmembrane protein gp41 from human
XX CC immunodeficiency virus 1 (HIV-1) isolate LA1. The DP107 peptide
XX CC corresponds to amino acids 558-595 of gp41 from HIV-1LA1. The invention
XX CC also relates to a method of identifying compounds that inhibit the
XX CC formation of or disrupts a DP107/DP178 complex. The method comprises
XX CC detecting the formation of a DP107/DP178 complex, both in the presence
XX CC and in the absence of a test compound, in a reaction mixture containing DP107
XX CC and DP178 peptides. The method is useful for identifying compounds,
XX CC including small molecule compounds, which may themselves exhibit
XX CC antifusogenic, antiviral or intracellular modulatory activity. The
XX CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
XX CC retroviral, particularly HIV, transmission to uninfected cells. The
XX CC present sequence represents a peptide sequence from Staphylococcus aureus
XX CC enterotoxin type E.
XX SQ Sequence 257 AA;

Query Match 31.4%; Score 1107; DB 22; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.1e-59;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Qy 226 SEKSEINEKDLRKSELOQTALGNLKOIYYYNKAITSSSEKSAQDQFLNTLLFKGFFTG 285
Db 25 SEKSEINEKDLRKSELOQTALGNLKOIYYYNKAITENKESDDQFLNTLLFKGFFTG 84

Qy 286 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 345
Db 85 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 144

Qy 346 BEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 405
Db 145 BEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204

Qy 406 RGLIVFHSSEGSTVSVDLFDAGGQVPTDQLLRIYRDN-----TISSTSLSL 458
Db 205 RGLIVFHSSEGSTVSVDLFDAGGQVPTDQLLRIYRDNKTINSENHLIDLYTINSENHLIDL 257

RESULT 4
AAG65572
ID AAG65572 standard; Protein; 720 AA.
XX AC AAG65572;
XX DT 30-NOV-2001 (first entry)
XX DE Amino acid sequence of a plasmid pFCAH9-E8d.
XX KW Gene library; immunoglobulin; antibody library.
XX OS Synthetic.
XX PN W0200162907-A1.
XX PD 30-AUG-2001.
XX PF 22-FEB-2001; 2001WO-JP01298.
XX PR 22-FEB-2000; 2000JP-0050543.
XX PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.
XX PI Kurosawa Y, Akahori Y, Iba Y, Morino K, Shinohara M, Takahashi M;
XX PI Okuno Y, Shiraki K;
XX WPI; 2001-565420/63.
XX DR N-PSDB; AAH47738.
XX PT Producing gene libraries and antibody libraries, involves selecting a
XX PT light chain that binds to a heavy chain product to produce a functional
XX PT formation, and producing a gene library of the light chain variable
XX PT regions -
XX PS Examples; Fig 3-6; 181pp; Japanese.
XX CC The invention relates to producing gene libraries, comprising
XX CC immunoglobulin light and heavy variable region. The method involves
XX CC selecting light chain that binds with the heavy chain product to produce
XX CC a functional conformation, producing a gene library comprising a
```

CC collection of these light chain variable genes, and combining with gene
CC library of heavy chain variable genes. The method is used for production
CC of gene and antibody libraries. The present sequence represents the
CC amino acid sequence of a plasmid pFCAH9-E8d.

```
XX SQ Sequence 720 AA;
Query Match 31.1%; Score 1096.5; DB 22; Length 720;
Best Local Similarity 37.9%; Pred. No. 9e-58;
Matches 256; Conservative 39; Mismatches 55; Indels 325; Gaps 9;
Qy 1 EVQLQSGDPLVKPGASVKISKASGYFTGYMHVWVKQSPGKLEWIGRINPNNGVTLY 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 QVQLQSGAELVKPGASVKLSCTASGPNIKDTYMHVWVKQRPKEG----- 66
Qy 61 NQKFKDKATLVDRKSTTAYMELSLTSEDSAVYICARSTMITNY---VMDYWGQGTSVT 117
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
67 -----LTSEDYAVYCA-----GYDYGNFYWGQGTTVT 95
Qy 118 VSSAKTPPSPVPLAPGSAOTNSMTLGCLVKGVPEPVTWNKSGLSGCVHTFPAVL 177
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 VSSASTKGPSVFPAPSPSKSTSGGTAALGCLVKDYFPEPVTWNKSGALTSGVHTFPAVL 155
Qy 178 QSD-LYTLSSSVTPSPSTWTPSETVTCNVAHPASSTKVDKIKIVPDSGGPSEKSEEINEKD 236
Db ||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 QSSGLYSLSSVTPSPSSSLGTOTVTCNVNHPKPSNTKVDKVEPK----- 199
Qy 237 LRKKELOGTALGNLKOIYYNKAITSSEKSAOFLTNTLLFKGFTGHPWYNDLLVDL 296
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 ----- 199
Qy 297 GSTAATSEYEGSSVDLYGAYGYOCAGCTPNKTCMYGGVTLHDNNRLTEKKVPINLWI 356
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 -----SCHH----- 203
Qy 357 DGKQTPVPIDKVTSSKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQVRLIVFHSSEG 416
Db :||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 -----HHHMKYLLPTAAA-----GLLL----- 221
Qy 417 STVSVDLFDAGQVPTLLRIYRDNNTTISSTLSLSLYLTSIVMTQTPSTLLVSAGDR 476
Db :||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 -----LAAQAPAMATSDIELTQSPASLSASVGET 249
Qy 477 VTITCKASQSVNDVAMVYQKPGQPKLLISYTSRYAGVPDRFSGSGYGTDFTLTISV 536
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
250 VTITCRASGNTHNYLA----- 265
Qy 537 QAEDAAYVFCQODYNSPPTFGGKTKLEIKRADAAPTYSIFPPSSQLTSGGASVVCFLNN 596
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 -----KLEIKRADAAPTYSIFPPSSQLTSGGASVVCFLNS 301
Qy 597 FYPKDIYKWKIDGSENGVLSNTDQSDKSTYSMSSTLTLPKDEYERHNSYTCEATH 656
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
302 FYPKDIYKWKIDGSENGVLSNTDQSDKSTYSMSSTLTLPKDEYERHNSYTCEATH 361
Qy 657 KTSTSPIVKSFNRNE 671
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
362 KTSTSPIVKSFNRNE 376
```

RESULT 5
AAB67339
ID AAB67339 standard; peptide; 230 AA.
XX
AC AAB67339;
XX
DT 23-APR-2001 (first entry)
XX
DE Staphylococcus aureus enterotoxin E protein.
XX
KW Tumour; cancer; immune; enterotoxin.
XX
OS Staphylococcus aureus.

```
XX US6180097-B1.  
PN  
XX  
PD 30-JAN-2001.  
XX  
PF 30-OCT-1998; 98US-0183437.  
XX  
PF 31-JAN-1994; 94US-0189424.  
PR 19-JUN-1995; 95US-0491746.  
PR 03-OCT-1989; 89US-0416530.  
PR 17-JAN-1990; 90US-0466577.  
PR 17-JAN-1991; 91WO-US00342.  
PR 01-JUN-1992; 92US-0891718.  
PR 02-MAR-1993; 93US-0025144.  
XX  
PA (TERM/) TERMAN D S.  
XX  
XX Terman DS;  
PI  
XX  
XX WPI; 2001-158657/16.  
DR  
XX  
XX Tumor cell capable of stimulating antitumor immune reactivity in vitro  
PT or in vivo comprises exogenous nucleic acids encoding a superantigen  
PT and a costimulatory molecule  
XX  
XX Disclosure; Fig 2; 16pp; English.  
PS  
XX  
XX The present invention relates to a tumour cell capable of stimulating  
CC antitumor immune reactivity in vitro or in vivo contains and  
CC expresses an exogenous nucleic acid molecule encoding a superantigen  
CC or its active fragment and an exogenous nucleic acid molecule  
CC encoding a costimulatory molecule that activates T cells in  
CC conjunction with an antigenic stimulus. The invention may be used  
CC for cancer therapy by stimulating an anticancer immune response  
CC in vivo or ex vivo.  
XX  
SQ Sequence 230 AA;  
Query Match 31.0%; Score 1093; DB 22; Length 230;  
Best Local Similarity 89.6%; Pred. No. 3.7e-58;  
Matches 206; Conservative 9; Mismatches 15; Indels 0; Gaps 0;  
Qy 229 SEENEKDLRKKSELOGTALGNLKOIYYNKAITSSEKSAOFLTNTLLFKGFTGHPW 288  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1  
1 SEENEKDLRKKSELOGTALGNLKOIYYNKAITSSEKSAOFLTNTLLFKGFTGHPW 60  
Qy 289 YNDLLVDLGSSTAATSEYEGSSVDLYGAYGYOCAGCTPNKTCMYGGVTLHDNNRLTEK 348  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 61  
61 YNDLLVDLGSSTAATSEYEGSSVDLYGAYGYOCAGCTPNKTCMYGGVTLHDNNRLTEK 120  
Qy 349 KVPINLWIDGKQTPVPIDKVTSSKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQVRL 408  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 121  
121 KVPINLWIDGKQTPVPIDKVTSSKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQVRL 180  
Qy 409 IVFHSSEGSTVSVDLFDAGQVPTLLRIYRDNNTTISSTLSLSLYLTT 458  
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 181  
181 IVFHSSEGSTVSVDLFDAGQVPTLLRIYRDNNTTISSTLSLSLYLTT 230
```

RESULT 6
AAW35374
ID AAW35374 standard; peptide; 245 AA.
XX
AC AAW35374;
XX
DT 20-APR-1998 (first entry)
XX
DE Staphylococcus enterotoxin SEE wild-type superantigen.
XX
KW SEE; SEA; Staphylococcus enterotoxin; superantigen; conjugate;
XX treatment; cancer; infection; autolimmune disease; antibody.
OS Staphylococcus sp.

FT	Misc-difference	20	"can be mutated at this position"
FT	Misc-difference	21	"can be mutated at this position"
FT	Misc-difference	24	"can be mutated at this position"
FT	Misc-difference	27	"can be mutated at this position"
FT	Misc-difference	27	"can be mutated at this position"

[illegible]

KW Single chain antibody; SCFv; inflammatory disease; arthritis; cancer;
KW hypersensitivity; autoimmune disease; central nervous system disorder;
KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
KW Helicobacter-related disease; immune disorder

BEST LOCAL SIMILARITY 37.96; FREQ. NO. 3.2E-30;

	Matches	254; Conservative	20; Mismatches	60; Indels	336; Gaps
QY	1	EVQLQQSGPDLVPGASVKISCKASGYFTGYHHWKQSPGKLEWIGRINPNNGVTLY	60		
Db	20	EVQLQQSGPDLVPGASVKISCKASGYFTGYHHWKQSHGKSLWIGRINPNNGVTLY	79		
QY	61	NQKFDKATLTVDKSSITAYMELRSLTSEDSAVYYCARSTMITNYVMYWGQGSVTVSS	120		
Db	80	NQKFDKAAILTVDKSSITAYMELRSLTSEDSAVYYCARSTMITNYVMYWGQGSVTVSS	138		
QY	121	AKTTPPSVYPLAPGSAQTNSMVTLGCLVKYFPEPTVTWNSGSLSGVHTFPFVQLQSD	180		
Db	139	-----	138		
QY	181	LYTLSSSVTPSPSTWPSSETVTCNVAHPASSTKVDKKIVPRDSGGPSEKSEINEKDLRKK	240		
Db	139	-----	141		
QY	241	SELQGTALGNLKOIYYNSKAITSSEKSAQDLTNTLLFKGFTGHPWYNLLVDLGGSTA	300		
Db	142	-----	141		
QY	301	ATSEYEGSSVDLYGAYGYQCAGCTPNKTCACMGVTLHDNNRLTEKKVPINLWIDGKQ	360		
Db	142	-----	151		
QY	361	TTVPIDKVKTSKKEVTVQELDLQARHYLHGFEGLYNSDSFGGKVGQGLIVFHSSEGSVTS	420		
Db	152	-----	153		
QY	421	YDLFDAOGYPTDLLRIYRDNTTISSTLSLSLYLTTSIVMTQTPTSLLYSAGDRVIT	480		
Db	154	-----	176		
QY	481	CKASQSVNDVAVYQQKPGQSPKLLISYTSRYAGVPRFSGSGYGTDTFTLTSSVQAE	540		
Db	177	CKASQSVNDVAVYQQKPGQSPPTLLISYTSRYAGVPRFICSGYGTDTFTLTSLQAE	236		
QY	541	AAVYFCQDYNSPPTFGGTKLEIKRADA-APTVSIPPPSEOLTSGCASVVCFLNNFVP	599		
Db	237	LAVYFCQDYNSPPTFGGTKLEIKRASKTGPSVFPPLAPSSKSTSGGTAALGCLVKDYFP	296		
QY	600	KDINWKK----IDGSEKQNGVLNSWTDQDSKSTYSMSSTLTLTCKDEYERHNSVTCBAT	655		
Db	297	EPVTVSNSGALTSQVHTFPFVQLQS-----SGLYSLSVTVTPVPSLSLGT-QTVICNVN	348		
QY	656	HKTSTPIVK 665			
Db	349	HKPSNTKVDK 358			
RESULT	8				
AAW86003	ID	AAW86003 standard; Protein; 595 AA.			
XX	XX				
AC	XX	AAW86003;			
XX	XX				
DT	XX	15-MAR-1999 (first entry)			
XX	XX				
DE	XX	Anti-5T4 single chain antibody 5T4sabl.			
KW	KW	Tumour interacting protein; cancer; gene therapy; vector;			
KW	KW	5T4 antigen; monoclonal antibody; single chain antibody;			
XX	XX	mouse; human; 5T4sabl.			
OS	XX	Chimeric - Mus sp.			
OS	XX	Chimeric - Homo sapiens.			
XX	XX	Chimeric - synthetic.			
PN	PN	W09855607-A2.			
XX	XX				
PD	XX	10-DEC-1998.			
XX	XX				

PF	04-JUN-1998;	98WO-GB01627.	
XX			
04	JUL-1997;	97GB-0014230.	
PR			
04	JUN-1997;	97GB-0011579.	
PR			
20	JUN-1997;	97GB-0013150.	
XX			
PA	(OXFO-)	OXFORD BIOMEDICA UK LTD.	
XX			
XX	Bebington CR,	Carroll MW, Ellard FM, Kingsman SM;	
PI	Myers KA;		
PI			
XX			
XX	WPI: 1999-059910/05.		
DR	N-PSDB; AAV80291.		
XX			
XX	New vector encoding a tumour interacting protein for treating cancer		
PT	- contains a desired nucleotide sequence and/or protein which		
PT	recognises tumours, and is used as a gene delivery system to treat		
PT	cancer		
XX			
PS	Example 1; Fig 1B; 82pp; English.		
XX			
CC	This is the amino acid sequence of a single chain antibody (sab),		
CC	termed 5T4sab1, comprising an scFv derived from murine monoclonal		
CC	antibody 5T4 (see AAW86002) and the human g1 constant region. cDNA		
CC	(see AAV80291) encoding the Sab has been inserted into vector pCIneo		
CC	to allow expression in mammalian cells. The trophoblast cell		
CC	surface antigen defined by 5T4 is expressed at high levels on the		
CC	cells of a wide variety of human tumours. The invention relates to		
CC	a vector comprising a nucleotide sequence coding for a tumour		
CC	interacting protein (TIP) and optionally a nucleotide sequence of		
CC	interest (NOI) which encodes a protein of interest (POI), the vector		
CC	being capable of delivering the NOI and/or POI to the tumour		
CC	recognised by the TIP. Delivery can be in vivo or ex vivo. The		
CC	vector is used to treat cancer, and may also used as a gene		
CC	delivery system for introducing at least 1 gene encoding a TIP		
CC	(preferably a tumour binding protein) into a haematopoietic cell		
CC	lineage.		
XX			
SQ	Sequence 595 AA:		
	Query Match	30.48; Score 1069; DB 20; Length 595;	
	Best Local Similarity	37.98; Pred. No. 3.2e-56;	
	Matches 254; Conservative	20; Mismatches 60; Indels 336; Gaps	
QY	1	EVQLQSQGPDLVKPGASVKISCKASGYSTFGYYMHVWQSPGKGLEWIGRINPNNGVTLY 60	
Db	23	EVQLQSQGPDLVKPGASVKISCKASGYSTFGYYMHVWQSPGKGLEWIGRINPNNGVTLY 82	
QY	61	NQKFQDKATLVDRKSTTAYMELRSLTSEDSAVYICARSTMTINYMDYWGOGTSVTYSS 12	
Db	83	NQKFQDKATLVDRKSTTAYMELRSLTSEDSAVYICARSTMTINYMDYWGOGTSVTYS- 14	
QY	121	AKTTPPSVYPLAGSAAQTNSMVTLGCLVKYGPPETVTWNSGSLSSGSHVHTFPAPVQLSD 18	
Db	142	----- 14	
QY	181	LYLTSSSVTPSSWPSETVTCNVAHPASSTKVDKKIVPRDGGPSEKSEEINEKDLRKK 24	
Db	142	-----SGC----- 14	
QY	241	SELQGTALGNLKQIYYNYSKAITSEKSAQOFLNTLLFKGFTGHPWYNDLLVLGSTA 30	
Db	145	----- 14	
QY	301	ATSEYEGSSVDLYGAYGYQCAGGTPPNKTAQMYGGVTLHDNNRLTEBKVPINLWDGKQ 36	
Db	145	-----CGSGGGCT----- 15	
QY	361	TTVPIDKVKTSKKEVTVOELDIQARHYLHCKFGLYNSDSFGKQVQKGLIVFHSSSGSTVS 42	
Db	155	-----GG----- 15	
QY	421	YDLFDAQOGYPDTLLRIYRDNTTTSSTLSISLYLTTSIVMTOTPTSLLVASAGRVTTIT 48	

XX WO9515382-A.
PN 08-JUN-1995.
XX 29-NOV-1994; 94WO-GB02610.
XX 03-JUN-1994; 94GB-0011089.
PR 03-DEC-1993; 93GB-0024819.
XX (ZENE) ZENECA LTD.
XX
XX Blakey DC, Boot C, Copley CG, Hall SM, Paterson DS;
PI Rose MS, Wright AF;
XX WPI: 1995-215262/28.
XX
XX Antigen binding structures containing CDRs recognising the CA55.1
PT antigen - produced by hybridomas and host cells, for use in the
PT diagnosis and therapy of cancer
XX
XX Claim 3; Page 97-98; 121pp: English.
XX
XX An antigen binding structure is based on the CDRs (given in AAR76078-
CC 84) of the heavy (AAR76085) and light (AAR76086) chains of Mab 55.1
CC (ECACC 93081901), which recognises the colorectal tumor-associated
CC antigen CA55.1. It is optionally humanized and in the form F(ab')₂,
CC F(ab)', Fab', Fv, scFv or v-min, and is produced in transgenic
CC animals or plasmids.
XX
XX Sequence 445 AA;
XX
Query Match 30.0%; Score 1055.5; DB 16; Length 445;
Best Local Similarity 37.5%; Pred. No. 1.5e-55;
Matches 250; Conservative 60; Mismatches 108; Indels 249; Gaps 14;
Qy 1 EVQLQQSPDLVLPKPGASVKISKASGYSTGYMHWKQSPGKLEWIGRINPNNGVTLY 60
Db 1 QVQLQPGAEVLVPGASVQLSCKASGYTFTGWIHWKQSPGKLEWIGVNPSTGRSDY 60
Qy 61 NQKFKRATLVDSSTTAYMELSLTSEDSAVYCARSTMI-TNYVMYDYGQGTSTVTS 119
Db 61 NEKFKRATLVDSSTTAYMQLSLSLTSEDSAVYCARERAYGYDDAMDYWGQGTSTVTS 120
Qy 120 SAKTTPSPVYPLAGSAQAOTNSMTVLGCLVKGPPEPTVTWNSGSLSSGVTTPAVLQS 179
Db 121 SAKTTPSPVYPLAGSAQAOTNSMTVLGCLVKGPPEPTVTWNSGSLSSGVTTPAVLQS 180
Qy 180 DLYTLSSSVTPSPSTWPTSETVTCNVAHPASSTKVDKIKVPRDSGSPSEKSEINEKDLRK 239
Db 181 DLYTLSSSVTPSPSTWPTSETVTCNVAHPASSTKVDKIKVPRDCG----- 224
Qy 240 KSELQGTALGNLKIYYNYSKAITSSSEKSAQDLNTLLFKGFTGHPWYNLLVDLGST 299
Db 225 -CKPCICTVPEVSSVFIFPPK-----PKDVLITL----- 253
Qy 300 AATSEYEGSSVDLYGAYGYQCAGGTGPNKTACMYGGVTLHDNRLTEKKVPINLWIDGK 359
Db 254 -----TP-KVTCVVVDIS-----KDDPEVQFNSFWFD-- 278
Qy 360 QTTVPIDVKVTSKREV-----TVQELDLQARHYLHGK-FCL-YNSDSFGKVGQGLIV 410
Db 279 -DVEVHTAQTPREEQFNSTFRSVSELPIMHQDLNGLNGKFKCRVNSAFAFPATIEK---- 332
Qy 411 FHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNTTISSTLSLSLYLYTTSIYMTQTPSLL 470
Db 333 -----TISKT----- 337
Qy 471 VSAGDRVITTCASQSVSNDAWYQKPGQSPKLLISYSSRYAGVDPFRFSGSGYGTDTF 530
Db 338 ----- 337
Qy 531 LTISVQAEADAAYFCQDYNPPTFGGGTKLEIKRADAPTIVSIFPPSSEQLTSGGASV 590

Db 338 -----KGRPKAPQVYTIPPPKEQAKDKVSL 363
Qy 591 VCFLNNFPKDVNVKWKIDGSRQGVNSWTQDQSKDSTYSMSSTLTLTDEYERHNSY 650
Db 364 TCMITDFPEDITVEMWONGPAEN-YKNTQPIMDT-DGSYFVYSKLVQKSNWEAGTF 421
Qy 651 TCEATHK 657
Db 422 TCSVLHE 428
RESULT 11
AAR45012
ID AAR45012 standard; protein; 230 AA.
XX AAR45012;
AC
XX 08-JUN-1994 (first entry)
XX Staphylococcal enterotoxin SEF.
XX
XX Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
KW autoimmune disease; toxicity; Protein A; perfusion system.
XX
XX Staphylococcus aureus.
XX
XX Key Location/Qualifiers
FT Misc-difference 120 /note= "Given in the specification as J, no further
FT details given"
FT Misc-difference 121 /note= "Given in the specification as J, no further
FT details given"
FT Misc-difference 123 /note= "Given in the specification as O, no further
FT details given"
FT Misc-difference 124 /note= "Given in the specification as U, no further
FT details given"
XX WO9324136-A.
PN
XX 09-DEC-1993.
PD
XX 01-JUN-1993; 93WO-US05213.
PF
XX 01-JUN-1992; 92US-0891718.
PR
XX (STON/) STONE J L.
PA (TERM/) TERMAN D S.
XX
XX Stone JL, Terman DS;
XX WPI: 1993-405418/50.
XX Use of staphylococcal enterotoxin(s) and homologues - for
PT treating cancer in a patient or for the treatment of auto-immune
PT diseases
XX
XX Disclosure; Fig 1; 90pp: English.
PS
XX The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
CC which may be used in the methods of the invention for treating cancer
CC in a patient. These SEs, and homologues of them, can be used as
CC tumouricidal agents for treating cancers and autoimmune disease.
CC They exhibit tumouricidal activity and toxicity identical to that
CC observed for the Protein A perfusion system. They may be administered
CC by i.v. injection.
XX
XX Sequence 230 AA;
SQ
Query Match 29.8%; Score 1048; DB 14; Length 230;

Best Local Similarity 85.7%; Pred. No. 1.9e-55;
Matches 197; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

Qy 229 SEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSAOQFLNTLLFKGFTGHPW 288
Db 1 SEEINEKDLRKSELORNALNLRQIYYNEKAITENKESDDQFLENTLLFKGFTGHPW 60

Qy 289 YNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKACMYGGVTLHDNNRLTEEK 348
Db 61 YNDLLVDKSGSDATNKYKGGKVDLYGAYGYOCAGGTPNKACMYGGVTLHDNNRLTEEX 120

Qy 349 KVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQRGL 408
Db 121 VYXXBKWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQRGL 180

Qy 409 IVFHSSEGSTVSYDLFDAQQGYPTDRLRIYRDNNTTISSTLSISLYLYTT 458
Db 181 IVFHSSEGSTVSYDLFDAQQGYPTDRLRIYRDNNTTISSENHMHIDIYLYTT 230

RESULT 12
ABB76235
ID ABB76235 standard; Protein: 230 AA.
AC ABB76235;
XX
DT 09-AUG-2002 (first entry)
XX
DE Staphylococcus aureus enterotoxin E.
XX
KW Enterotoxin E; SEE; superantigen; antigen; tumour; cancer;
KW antitumour; therapy.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 120
FT /note= "given as 'J' in the specification"
FT Misc-difference 121
FT /note= "given as 'J' in the specification"
FT Misc-difference 123
FT /note= "given as 'O' in the specification"
FT Misc-difference 124
FT /note= "given as 'U' in the specification"
FT Misc-difference 125
FT /note= "given as 'V' in the specification"
XX
PN US2002051765-A1.
XX
PD 02-MAY-2002.
XX
PF 19-DEC-2000; 2000US-0741503.
XX
PR 31-JAN-1994; 94US-0189424.
PR 19-JUN-1995; 95US-0491746.
PR 03-OCT-1989; 89US-0416530.
PR 17-JAN-1990; 90US-0466577.
PR 17-JAN-1991; 91WO-US00342.
PR 01-JUN-1992; 92US-0891718.
PR 02-MAR-1993; 93US-0025144.
XX
PA (TERM/) TERMAN D S.
XX
XX Terman DS;
XX
XX WPI; 2002-415198/44.
XX
XX Reagent for treating cancer without the need for e.g. radiotherapy,
XX comprises a specific V beta subset of T cells sensitized to a growing
XX tumor and stimulated with superantigens
XX
XX Disclosure; Fig 2; 17pp; English.
XX

CC The present sequence is the protein sequence of enterotoxin E
CC (SEE) of Staphylococcus aureus. Similarity is shown, in several
CC stretches of sequence, between staphylococcal enterotoxins,
CC streptococcal pyrogenic exotoxins and staphylococcal exfoliative
CC toxins (see ABB76234-44). In the present invention, synthetic
CC polypeptides useful in tumour therapy and in blocking or destroying
CC autoreactive T and B lymphocyte populations are characterised by
CC substantial structural homology to staphylococcal enterotoxin A and
CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with
CC statistically significant sequence homology and similarity (% value
CC of Liman and Pearson algorithm in Monte Carlo analysis exceeding
CC 6) to include alignment of cysteine residues and similar tumour
CC profiles. These superantigens are used to treat solid tumours,
CC including their metastases, without radiation, surgery or
CC standard chemotherapeutic agents. A claimed method of human cancer
CC treatment involves contacting haematopoietic cells from a patient
CC with one or more superantigens ex vivo to generate stimulated cells,
CC selecting a specific V beta subset of cells, and reintroducing
CC these cells into the patient to induce an in vivo therapeutic,
CC tumouricidal reaction.
XX
SQ Sequence 230 AA;
Query Match 29.6%; Score 1044; DB 23; Length 230;
Best Local Similarity 85.7%; Pred. No. 3.2e-55;
Matches 197; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

Qy 229 SEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSAOQFLNTLLFKGFTGHPW 288
Db 1 SEEINEKDLRKSELORNALNLRQIYYNEKAITENKESDDQFLENTLLFKGFTGHPW 60

Qy 289 YNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKACMYGGVTLHDNNRLTEEK 348
Db 61 YNDLLVDKSGSDATNKYKGGKVDLYGAYGYOCAGGTPNKACMYGGVTLHDNNRLTEEX 120

Qy 349 KVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQRGL 408
Db 121 VYXXBKWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQRGL 180

Qy 409 IVFHSSEGSTVSYDLFDAQQGYPTDRLRIYRDNNTTISSTLSISLYLYTT 458
Db 181 IVFHSSEGSTVSYDLFDAQQGYPTDRLRIYRDNNTTISENHMHIDIYLYTT 230

RESULT 13
AAR13204
ID AAR13204 standard; Protein: 230 AA.
XX
AC AAR13204;
XX
DT 15-OCT-1991 (first entry)
XX
DE Staphylococcal enterotoxin E.
XX
KW SEE; cancer treatment; pyrogen; tumouricide.
XX
OS Staphylococcus aureus.
XX
PN WO9110680-A.
XX
PD 25-JUL-1991.
XX
PF 17-JAN-1991; 91WO-US00342.
XX
PR 17-JAN-1990; 90US-0466577.
XX
PA (TERM/) TERMAN D S.
XX
XX Terman DS;
XX
XX WPI; 1991-237984/32.
XX
PT Treating cancer with enterotoxin from Staphylococcus aureus -

PT administered by IV injection, having same tumoricidal activity
 PT as Staphylococcal protein A without potential toxic reactions

PS Disclosure; Fig 1; 74pp; English.

XX SEE was isolated and purified from S.aureus. It can be used for
 CC treating cancer, activating cytokine mediators and procoagulant
 CC systems, augmenting natural killer cell cytotoxicity, etc. The
 CC enterotoxin can be administered intravenously, optionally with
 CC ibuprofen to attenuate toxic reaction to SEE. Synthetic
 CC polypeptides having structural homology to Staphylococcal exotoxins
 CC are claimed, provided the homology includes statistically
 CC significant sequence homology, alignment of cysteine residues and
 CC similar hydrophathy profiles.

XX Sequence 230 AA;

Query Match 29.4%; Score 1035; DB 12; Length 230;
 Best Local Similarity 84.3%; Pred. No. 1.1e-54;
 Matches 194; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

Qy 229 SEETNEKDLRKKSELOQTALGNLQIYYNKAITSSEKSDQFLTNLLPKGFTGHPW 288

Db 1 SEETNEKDLRKKSELOQTALGNLQIYYNKAITSSEKSDQFLTNLLPKGFTGHPW 60

Qy 289 YNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTTEK 348

Db 61 YNDLLVDLGSGDANKYKGVKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTTEK 120

Qy 349 KVPINLWIDGKQTTPVIDKVTSSKEVTQVELDLQARHYLHGKFLYNSDSFGGKVORGL 408

Db 121 XVQXKWDIGKQTTPVIDKVTSSKEVTQVELDLQARHYLHGKFLYNSDSFGGKVORGL 180

Qy 409 IVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNNTISSTLSISLYLTT 458

Db 181 IVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNNTISSTLSISLYLTT 230

RESULT 14

ID AE18377 standard; Protein: 597 AA.

XX AE18377;

XX 07-MAY-2002 (first entry)

DE Human N-terminal DAV-1 heavy chain-mature TNF-alpha fusion protein.

XX Human; bifunctional molecule; monoclonal antibody; gene therapy; cancer;
 KW vascular disorder; diabetic retinopathy; restenosis; ophthalmic disorder;
 KW hyperproliferative disorder; hormonal disorder; DAV-1 heavy chain;
 KW cytosstatic; vasotropic; ophthalmological; tumour necrosis factor-alpha;
 KW TNF-alpha; fusion protein.

OS Chimeric - Homo sapiens.

OS Chimeric - Synthetic

XX Key Location/Qualifiers

FT Region 1..439

FT /note= "N-terminal portion of DAV-1 heavy chain"

FT Region 441..597

FT /note= "Human mature TNF-alpha"

XX WO2000204522-A2.

XX 17-JAN-2002.

XX 09-JUL-2001; 2001WO-Ep07878.

XX 10-JUL-2000; 2000US-0613017.

XX (NOVS) NOVARTIS AG.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 PA (SCRI) SCRIPPS RES INST.

XX Nemerow GR, Li E;

XX WPI; 2002-171707/22.

XX New bifunctional molecules comprising an antibody or its
 PT antigen-binding portion, and a targeting agent, useful for e.g. gene
 PT therapy, or for promoting Adenoviral vector-mediated gene delivery to
 PT cells lacking av integrins

XX Claim 15; Page 98-99; 106pp; English.

XX The present invention relates to a bifunctional molecule comprising an
 CC antibody or its antigen-binding portion, and a targeting agent where the
 CC antibody specifically binds to an antigen in a protein that binds to av
 CC integrin, and the targeting agent specifically binds to a cell surface
 CC protein that activates the phosphatidylinositol 3 (PI3K) signalling
 CC pathway. The bifunctional molecules are useful for gene therapy, for
 CC promoting Adenoviral (Ad) vector-mediated gene delivery to cells lacking
 CC av integrins, for enhancing Ad binding and internalisation, and in gene
 CC delivery of by fibreless adenovirus particles. The bifunctional molecules
 CC permit targeting of viral and bacterial vectors to cells that cancers,
 CC targeted receptors. Diseases that can be targeted include cancer,
 CC vascular disorders, diabetic retinopathies, restenosis, ophthalmic
 CC disorders, hyperproliferative disorders, and hormonal disorders.
 CC The present sequence is human N-terminal DAV-1 heavy chain-mature tumour
 CC necrosis factor-alpha (TNF-alpha) fusion protein which is used in the
 CC invention.

XX Sequence 597 AA;

Query Match 29.2%; Score 1028; DB 23; Length 597;
 Best Local Similarity 37.2%; Pred. No. 9.4e-54;
 Matches 253; Conservative 53; Mismatches 115; Indels 260; Gaps 15;

Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSTFTGYMHVWVKQSPGKLEWIGRINPNNGVTLY 60

Db 20 EVQLQQSGPELVKPGASVKISKASGYSTFTDYNMHVWVKQSHGKSLIEWIGYIPYKGGTGY 79

Qy 61 NQKFKDRLTLVDKSSVTAYMELRSLTSEDSAVVYCARSTMTNVMYDWGQGSVTVTSS 120

Db 80 NQKFKSKATLTDTSSNTAYMELRSLTSDASAVVYCARG-----IAYWGQGLVTVYSA 132

Qy 121 AKTTPPSVYPLAPGSAATNSMVTGLCLVKGYFPEPVTVTWNSGSLSGVHTFPAVLQSD 180

Db 133 AKTTPPSVYPLAPGSAATNSMVTGLCLVKGYFPEPVTVTWNSGSLSGVHTFPAVLQSD 192

Qy 181 LYTSSSVTPSPSTWSPSETVTCNVAHPASSTKVDKKIIPRDSGGPSEKSEEINEKDLRKK 240

Db 193 LYTSSSVTPSPSTWSPSETVTCNVAHPASSTKVDKKIIPRDCG----- 235

Qy 241 SELOGTALGNLKOIYYNKAITSSEKSDQFLTNLLFKGFTGHTGHPWYNDLLVDLGSTA 300

Db 236 CKPCICTVPEVSSVFIFPPK-----PKDVLTLT----- 264

Qy 301 ATSEVEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTTEKKVPINLWIDCKQ 360

Db 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWEVD--- 289

Qy 361 TTVPIDKVKTSKKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVQRLIVF 411

Db 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHQDWLNGREFKCRVNSAAPPAPIEK----- 343

Qy 412 HSEGSTVSVDLFDAGQGYPTDLLRIYRDNNTISSTLSISLYLTTISIVMTQTPTSLV 471

Db 344 -----TISKT----- 348

Qy 472 SAGDRVTITCKASQSVNDVAWYQKPGOSPKLLISYTSRYSAGVYVDRFSGSGYCTDFTL 531

Db 349 ----- 348

QY 532 TISSVQAEADAAVYFCQDDYNSPPTFGGKTKLEIKRADAAPTVSIFFPPSSEQLTSGGASVY 591
Db 349 -----KGRKAPQVYTIPTPPKPEQMAKDKVSLT 375
QY 592 CFLNFFPKDINVKIKDGSERQGNVLSWTQDQSKDSTYSMSSTLTLTCKDEYERHNSYT 651
Db 376 CMITOFFPEDITVEQWNGOPAEV-YKNTQPTMDT-DGSFYVYSKLVNOKSNWEAGNTFI 433
QY 652 CEATHKTSPIVKSFNRNES 672
Db 434 CSVLHE-----FVRSSSRTPS 449
RESULT 15
AAW83041
ID AAW83041 standard; Protein; 464 AA.
XX AAW83041;
AC AAW83041;
XX 15-MAR-1999 (first entry)
DT Anti-Fas MAB HFE7A heavy chain.
DE
XX HFE7A; monoclonal antibody; mouse; Fas; humanised antibody;
KW apoptosis; HFE7A; autoimmune disease; Hashimoto's disease;
KW systemic lupus erythematosus; graft versus host disease;
KW Sjogren syndrome; pernicious anaemia; Addison's disease;
KW scleroderma; Goodpasture syndrome; Crohn's disease; sterility;
KW rheumatoid arthritis; autoimmune haemolytic anaemia;
KW myasthenia gravis; multiple sclerosis; Basedow's disease;
KW thrombopenia purpura; insulin-dependent diabetes; allergy;
KW atopy; arteriosclerosis; myocarditis; cardiomyopathy;
KW glomerular nephritis; hypoplastic anaemia; hepatitis; AIDS;
KW transplant rejection; therapy; complementarity determining region;
CDR.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= Sig_peptide
FT Protein 20..464
FT /label= Mat_protein
FT Region 20..140
FT /label= Variable
FT Region 141..464
FT /label= Constant
FT Region 50..54
FT /label= CDR_H1
FT /note= "claim 9"
FT Region 69..84
FT /label= CDR_H2
FT /note= "claim 9"
FT Region 118..128
FT /label= CDR_H3
FT /note= "claim 9"
XX
PN AU9859701-A.
XX
XX
PD 08-OCT-1998.
XX
PF 30-MAR-1998; 98AU-0059701.
XX
PR 08-OCT-1997; 97JP-0276064.
PR 01-APR-1997; 97JP-0082953.
PR 25-JUN-1997; 97JP-0169088.
XX
PA (SANY) SANKYO CO LTD.
XX
PI Akio S, Hideyuki H, Hiroko Y, Jun O, Kimihisa I;
PI Masahiko O, Nobufusa S, Shin Y, Tohru T;
XX
DR WPI: 1998-543440/47.

DR N-PSDB: AAV71029.
XX New antibodies and proteins bind conserved epitope of Fas antigen -
PT used to evaluate drugs in animal models and to treat Fas-associated
PT diseases e.g. autoimmune disease, allergy, atopy, arteriosclerosis,
PT myocarditis, hepatitis and AIDS
XX
PS Reference Example 4; Page 187-188; 292pp: English.
XX
XX This is the amino acid of the heavy chain of murine anti-human Fas
CC monoclonal antibody HFE7A. cDNA (see AAV70129) encoding the heavy
CC chain was obtained from HFE7A-secreting hybridoma (FERM Bp-5828)
CC RNA by RT-PCR (see AAV70125-26). The invention provides humanised
CC HFE7A antibodies (see AAW83031-37) produced by CDR grafting. These
CC antibodies are capable of inducing apoptosis in abnormal cells
CC expressing Fas, and of inhibiting Fas-induced apoptosis in normal
CC cells. They are used to evaluate, in animal models, treatments of
CC diseases that involve Fas/Fas ligand interactions, and also to treat
CC such diseases, including autoimmune disease (e.g. systemic lupus
CC erythematosus, Hashimoto's disease, graft versus host disease,
CC Sjogren syndrome, pernicious anaemia, Addison's disease,
CC scleroderma, Goodpasture syndrome, Crohn's disease, rheumatoid
CC arthritis, autoimmune haemolytic anaemia, sterility, myasthenia
CC gravis, multiple sclerosis, Basedow's disease, thrombopenia purpura
CC and insulin-dependent diabetes), allergies, atopy, arteriosclerosis,
CC myocarditis, cardiomyopathy, glomerular nephritis, hypoplastic
CC anaemia, hepatitis, AIDS and transplant rejection (all claimed).
XX
SQ Sequence 464 AA;
Query Match 29.1%; Score 1025.5; DB 19; Length 464;
Best Local Similarity 36.2%; Pred. No. 9.8e-54;
Matches 243; Conservative 64; Mismatches 106; Indels 259; Gaps 14;
QY 1 EVQLQSGPDIVKPGASVKISKASGYSTGYMHVWVQSPCKGLEWIGRINPNNGVTLY 60
Db 20 QVQLQPGAEIVKPGASVKLSCKAGYFTFTYMMQWVKRPGQGLEWIGEDIPDSYTN 79
QY 61 NQKFKDKATLTVDKSTTAYMELSLTSDSAVYVCARS-TWITNYVMYDQGSVTVS 119
Db 80 NQKFKGKATLTVDTSSTAYMQLSSLTSDSAVYVCARNRYSNNWYFDVWGTTVTVS 139
QY 120 SAKTTPPSVYPLAPGSAQTNSMVTLCGLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQS 179
Db 140 SAKTTPPSVYPLAPGSAQTNSMVTLCGLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQS 199
QY 180 DLYTLSSSVTPSPSTWPSQVTCNVNHPASSTKVDKIVPRDSGGPSEKSEINEKDLRK 239
Db 200 DLYTLSSSVTPSPSTWPSQVTCNVNHPASSTKVDKIVPRDCG----- 243
QY 240 KSELQGTALGNLKOIYYNYSKAITSEKSAQDQFLNTLLFKGFTGHPWYNLLVDLGST 299
Db 244 -CKPCICIVPEVSSVFIPPK-----PKDVLITL----- 272
QY 300 AATSEYEGSSVDLYGAYGYQCAGGTPNKACMYGGVTLHDNNRLTEKKVPINLWIDGK 359
Db 273 -----TP-KVTCVWVDIS-----KDDPEVQFSNFDV-- 297
QY 360 QTTVPIDKVKTSKKEV-----TVQELDLQARHLHGK-FGL-YNSDSFGKVGRLIV 410
Db 298 --DEVHTAQTPREEQFNSTERSVSELPIMHQNWLNGKEFKRCVNSAFAPIEK----- 351
QY 411 FHSSEGSTVSYDLFDAQGYPDTLRLIYRDNTTISTSLSLSLYLVTTSIVMTQTPTSL 470
Db 352 -----TISKT----- 356
QY 471 VSAGDRVTITCKASVSNDAVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYTDFT 530
Db 357 ----- 356
QY 531 LTISVQAEADAAVYFCQDDYNSPPTFGGKTKLEIKRADAAPTVSIFFPPSSEQLTSGGASV 590
Db 357 -----KGRKAPQVYTIPTPPKPEQMAKDKVSL 382

PF 29-SEP-1999; 99EP-0307711.
XX
PR 30-SEP-1998; 98JP-0276881.
PR 30-SEP-1998; 98JP-0276882.
XX
PA (SANY) SANKYO CO LTD.
XX
PI Serizawa N, Haruyama H, Nakahara K, Tamaki I, Takahashi T;
XX
DR WPI: 2000-258930/23.
DR N-PSDB; AA11546.
XX
XX New humanized anti-Fas antibody, useful for treating or preventing e.g.
PT inflammatory or autoimmune disease, induces apoptosis selectively in
PT cells with abnormal Fas-Fas ligand systems
XX
XX Example reference 4; Page 100-102; 263pp; English.
XX
CC This invention describes a novel humanized anti-Fas antibody-like
CC molecule (I) that, induces apoptosis in cells with an abnormal Fas/Fas
CC ligand system, by binding to Fas on the cell surface, and prevents
CC apoptosis in cells with a normal system, by inhibiting binding between
CC Fas and its ligand. The products of the invention have anti-inflammatory,
CC anti-anemic, antidiabetic, anti-allergic, anti-arthritis, antiviral,
CC immunomodulatory, dermatological, immunosuppressive, thymomimetic,
CC antirheumatic, nephrotropic, antiinfertility, neuroprotective,
CC antiarteriosclerotic, cardiant and hepatropic activity. (I) induce
CC apoptosis by binding to cell surface Fas or inhibit it by competitive
CC inhibition of ligand binding. (II) are used to treat and/or prevent
CC diseases associated with the Fas/Fas ligand system, especially systemic
CC lupus erythematosus, Hashimoto disease, rheumatoid arthritis, graft
CC versus host disease, Sjorgen's syndrome, pernicious or hypoplastic
CC anemia, Addison's disease, scleroderma, Goodpasture syndrome, Crohn's
CC disease, autoimmune hemolytic anemia, sterility, myasthenia gravis,
CC multiple sclerosis, Basedow's disease, thrombopenia purpura, insulin
CC dependent diabetes mellitus, allergy, arteriosclerosis, myocarditis,
CC cardiomyopathy, glomerulonephritis, hepatitis (fulminant, chronic, viral
CC (B, C or D) or alcoholic), and transplant rejection. (I) selectively
CC inhibit apoptosis in normal cells but selectively induce it in abnormal
CC cells. They bind to both human and murine Fas, so can be evaluated in
CC murine disease models. (I) act on the active site of Fas, i.e. they mimic
CC the native ligand, do not induce liver disease, and have reduced risk of
CC inducing a human anti-murine antibody response. This sequence represents
CC a murine anti-Fas monoclonal antibody HFE7A heavy chain described in the
CC method of the invention.
XX
SQ Sequence 464 AA;

Query Match 29.1%; Score 1025.5; DB 21; Length 464;
Best Local Similarity 36.2%; Pred. No. 9.8e-54;
Matches 243; Conservative 64; Mismatches 106; Indels 259; Gaps 14;

QY 1 EVQLQQSGDPLVKPGASVKISCKASGYSTGYVMHWKVGSPGKGLWIGRIINPNNGVTLY 60
DB :||||| : :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
20 QVQLQPGGAEIVKPGASVKLSKASGTTSTYMQWVKORPGOGLEWIGIDPDSSTYTN 79
QY 61 NQKFKKATLVTDKSTTAYMELRSLTSDSVAIVYCARSS-TMITNYVMYWGOGTSVTVS 119
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
80 NQKFKKATLVTDSSSTAYMQLSLTSDSVAIVYCARRDYSNNWYFDVWGCTGTTVTS 139
QY 120 SAKTTPSVYPLAPGSAATNSMTVLCGLVKGYFPPEVTVTWNSGSLSGVHTFPVAVLS 179
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
140 SAKTTPSVYPLAPGSAATNSMTVLCGLVKGYFPPEVTVTWNSGSLSGVHTFPVAVLS 199
QY 180 DLYTLSSSVTVPSSTWPSVTVCNVAHPASSTKVDKKIVPRDSGGSPSEKSEENEDLRK 239
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
200 DLYTLSSSVTVPSSTWPSVTVCNVAHPASSTKVDKKIVPRDCG----- 243
QY 240 KSELOGTALCNLKOIYYXNKAITSSEKSAQDFLTNTLLFKGFETHGHPWNLLVLGSLT 299
DB :
244 -CKPCLCTVPEVSSVFIFFPK-----PKDVLTTTL----- 272
QY 300 AATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEPKKVPINLWIDGK 359

Db 273 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD-- 297
QY 360 QTTVPIDKVKTSKEV-----TVEQLDQARHYLVHGR-FGL-VNSDSFGKGKVGRLIV 410
Db 298 --DVEVHTAQTPREERQFNSTFRSVSELDPMHQNWLNGKEFKCRVNSAAPPAPIEK----- 351
QY 411 FHSSEGSTVSYDLFDAQGGYPTDLRLIYRDNTTISTSTLSLSLYLYTTSIVMTQTPSTLL 470
Db 352 -----TISKT----- 356
QY 471 VSAGDRVTTTCASQSVSNDAVYQOKPCQSPKLLISYTSRSRYAGVDPDRFSGSGYGTDTF 530
Db 357 ----- 356
QY 531 LTISSVQAEDAAYVFCQDYNPPTFGGOTKLEIKRADAAPTYSIFPPSSSEOLTSGGASV 590
Db 357 -----KGRKAPQVYTIPTPKQMAKDKVSL 382
QY 591 VGFLNNFYPKDINVWKIDGSRQN-----GVLNSWTDQDSKDSYMSSTLTLTKEYE 645
Db 383 TCMITDFFPEDITVEQWNGQPAENYKNTQPIMNT-----NGSYFVYKLVQKSNWE 435
QY 646 RNSYTCETHAK 657
Db 436 AGNTFTCSVLHE 447
RESULT 18
ABB74866
ID ABB74866 standard; Peptide; 464 AA.
XX
AC ABB74866;
XX
DT 26-APR-2002 (first entry)
XX
DE Humanised anti-Fas antibody related peptide SEQ ID NO 32.
XX
KW Human; mouse; humanised anti-Fas antibody; Fas/Fas ligand;
KW light chain subunit; apoptosis; immunosuppressive; antiallergic;
KW autoimmune disease; allergy; atopic.
XX
OS Homo sapiens.
XX
PN JP2001342148-A.
XX
PD 11-DEC-2001.
XX
PF 28-MAR-2001; 2001JP-0093106.
XX
PR 29-MAR-2000; 2000JP-0090918.
XX
PA (SANY) SANKYO CO LTD.
XX
DR WPI; 2002-145113/19.
XX
PT Drug containing humanised anti-Fas antibody, used for preventing and
XX treating autoimmune diseases, allergy, and atopy -
XX
PS Example 6 (Preparatory); Page 26; 194pp; Japanese.
XX
CC The invention relates to a preventive or treating agent for diseases
CC caused by abnormality in Fas/Fas ligand system containing as the active
CC component an antibody having as the light chain subunit a polypeptide
CC containing residues 1-218 of one of 3, 239 residue amino acid sequences,
CC or residues 1-451 of one of 3, 470 residue amino acid sequences, all
CC fully defined in the specification and having an activity of combining
CC specifically with mammalian Fas and an activity of inducing apoptosis
CC in a cell expressing Fas. The agent has immunosuppressive and
CC antiallergic activity and is used for preventing and treating autoimmune
CC diseases, allergy, atopy and others. The present sequence is that of a
CC peptide, useful to the invention.
XX

FH	Key	Location/Qualifiers	
FT	Region	230...242	
XX		/note= "Hinge region"	
PN		WO200204522-A2.	
PD		17-JAN-2002.	
XX		09-JUL-2001; 2001WO-EP07878.	
XX		10-JUL-2000; 2000US-0613017.	
XX		(NOVS) NOVARTIS AG.	
PA		(NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.	
PA		(SCRI) SCRIPPS RES INST.	
XX		Nemerow GR, Li E;	
PI		WPI; 2002-171707/22.	
DR		N-PSDB; AAD29308.	
XX			
PT		New bifunctional molecules comprising an antibody or its	
PT		antigen-binding portion, and a targeting agent, useful for e.g. gene	
PT		therapy, or for promoting Adenoviral vector-mediated gene delivery to	
PT		cells lacking av integrins	
XX			
PS		Claim 10; Page 91-92; 106pp; English.	
XX			
CC		The present invention relates to a bifunctional molecule comprising an	
CC		antibody or its antigen-binding portion, and a targeting agent where the	
CC		antibody specifically binds to an antigen in a protein that binds to av	
CC		integrin, and the targeting agent specifically binds to a cell surface	
CC		protein that activates the phosphatidylinositol 3 (PI3K) signalling	
CC		pathway. The bifunctional molecules are useful for gene therapy, for	
CC		promoting Adenoviral (Ad) vector-mediated gene delivery to cells lacking	
CC		av integrins, for enhancing Ad binding and internalisation, and in gene	
CC		delivery of by fibreless adenovirus particles. The bifunctional molecules	
CC		permit targeting of viral and bacterial vectors to cells that express	
CC		targeted receptors. Diseases that can be targeted include cancers,	
CC		vascular disorders, diabetic retinopathies, restenosis, ophthalmic	
CC		disorders, hyperproliferative disorders, and hormonal disorders.	
CC		The present sequence is human penton base monoclonal antibody, DAV-1	
CC		heavy chain.	
XX			
SQ	Sequence	456 AA;	
	Query Match	29.1%; Score 1023.5; DB 23; Length 456;	
	Best Local Similarity	37.4%; Pred. No. 1.3e-53;	
	Matches 249; Conservative	51; Mismatches 111; Indels 255; Gaps 14;	
QY	1	EVQLQSGPDLVKPGASVKISCKASGYSTGYIMHWVKSPCKGLEWIGRINPNNGVTLY 60	
Db	20	EVQLQSGPDLVKPGASVKISCKASGYSTGYIMHWVKSHGKSLWIGIYIPYKGGTGY 79	
QY	61	NQKFKDKATLTVDKSTTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGQGTSTVSS 120	
Db	80	NQKFKSKATLTDTSSNTAYMELRSLTSDSAVYYCARG-----IAYWGQGLTIVTSA 132	
QY	121	AKTTPPSVYPLAPGSAQAQNSWVTLGCLIVKGYFPEPVTVWNSGSLSSGVHFTFAVLQSD 180	
Db	133	AKTTPPSVYPLAPGSAQAQNSWVTLGCLIVKGYFPEPVTVWNSGSLSSGVHFTFAVLQSD 192	
QY	181	LYTLSSSVTPSPSTWPEVTNCVAHPASSTKVDKIVPRDSGGPSEKSEINEKDLRKK 240	
Db	193	LYTLSSSVTPSPSTWPEVTNCVAHPASSTKVDKIVPRDCG----- 235	
QY	241	SELQGTALGNLKOIYYNNSKAITSEKSAQDFLTNTLFKGFETGHPWYNLLDLVGSTA 300	
Db	236	CKPCICTVEVSSVFIFPPK-----PKDVLITL----- 264	
QY	301	ATSEYEGSSVDLYGAYGYQCAGGTGPNKTACMYGGVTLHDNNRLTEEEKVPINLWIDGKQ 360	
Db	265	-----TP-KVTCVVVDIS-----KDDPEVQFSEFVD--- 289	

QY	361	TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGKVGQGLIVE 411	
Db	290	-DVEVHTAQTPRECFNSTFRSVSELPMTMHODWLNKFEKCRVNSAAPAIEK----- 343	
QY	412	HSSEGSTVSYDLFDAQGOYPTLLRIYRDNTTISSTLSISLYLYTTSIVMTQTPTSLIV 471	
Db	344	-----TISKT----- 348	
QY	472	SAGDRVTITCKASQSVSNVDVANYQKPGQSPKLLISYTSRIYAGVDPDRFSGSGYGTDFTL 531	
Db	349	----- 348	
QY	532	TISSVQAEAAVYFCQDYNSPPTGGGPKLEIKRADAAPTVSIRPPSSEQLTSGGASVV 591	
Db	349	-----KGRPKAPQVYTIPTPPKEQMAKDKVSLT 375	
QY	592	CFLNMFYPKDINVWKIDGSRQGNVLSWTDQDSKDYSSMSTLTLTQDEYERHNSYT 651	
Db	376	CMITDFPEDITVEQWNGQPAEN-YKNTPIMDT-DGSYFYVSKLVNOKSNWEAGNTFI 433	
QY	652	CEATHK 657	
Db	434	CSVLHE 439	

RESULT 22

AAE18379	standard; Protein; 493 AA.
XX	AAE18379;
DT	07-MAY-2002 (first entry)
DE	Human N-terminal DAV-1 heavy chain-mature EGF fusion protein.
XX	
KW	Human; bifunctional molecule; monoclonal antibody; gene therapy; cancer;
KW	vascular disorder; diabetic retinopathy; restenosis; ophthalmic disorder;
KW	hyperproliferative disorder; hormonal disorder; DAV-1 heavy chain;
KW	cytostatic; vasotropic; ophthalmological; epidermal growth factor;
XX	EGF; fusion protein.
OS	Chimeric - Homo sapiens.
OS	Chimeric - Synthetic
XX	
FH	Key
FT	Region
FT	Region
FT	Region
FT	Region
XX	WO200204522-A2.
XX	17-JAN-2002.
XX	09-JUL-2001; 2001WO-EP07878.
XX	10-JUL-2000; 2000US-0613017.
XX	(NOVS) NOVARTIS AG.
PA	(NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PA	(SCRI) SCRIPPS RES INST.
XX	Nemerow GR, Li E;
XX	WPI; 2002-171707/22.
XX	
PT	New bifunctional molecules comprising an antibody or its
PT	antigen-binding portion, and a targeting agent, useful for e.g. gene
PT	therapy, or for promoting Adenoviral vector-mediated gene delivery to
PT	cells lacking av integrins
XX	
PS	Claim 15; Page 101-102; 106pp; English.


```
QY 61 NQFKDKATLTVDKSTTAYMELRLTSDSAVYVCARSTMTNVMYDYGQGTSTVSS 120
PI ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 NQFKSKATLTDTSSNTAYMELRLTSDSAVYVCARG-----IAYWGGTLTVSA 132

QY 121 AKTTPSVYPLAPGSAQAQNSWVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 133 AKTTPSVYPLAPGSAQAQNSWVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 192

QY 181 LYTSSSVTPSPSTWPEVTCNVAHPASSTKVDDKIVPRDGGPSEKSEENKDLRKK 240
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 193 LYTSSSVTPSPSTWPEVTCNVAHPASSTKVDDKIVPRDCG----- 235

QY 241 SELQGTALGNLKOIYYNYSKAITSEKSDAQFLTNTLLFKGFFTGHPWYNDLLVLDGSTA 300
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 236 CKPCICTVEVSSVFIFPPK-----PKDVLITL----- 264

QY 301 ATSEYEGSSVDLYGAYGYQCAGGTGPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKO 360
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWEVD--- 289

QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVORGLIVF 411
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHODWLNKKEKCRVNSAAPPAPIEK----- 343

QY 412 HSEGSTVSVDLFDAGQGYPDTLRLIYRDNTTISLSISLYLYTTSIVMTQTPTSLLV 471
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 344 -----TISK-T----- 348

QY 472 SAGDRVTITCKASQSVNSDVAMVYQKPGSQKLLISYTSRYAGVPDRFSGSGYGTDFTL 531
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 349 ----- 348

QY 532 TISSVQAEDAAYVFCQDYNSPPTFGGKTKLEIKRADAAPTYSIRPPSSQELTSGGASVY 591
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 349 -----KGRPKAPQVITTPPPKQEMAKOKVSLT 375

QY 592 CFLNNFYPKIDINVKWKIDGSRQNGVLNSWTDQDSKDYSTMSSTLTLRKDEYERHNSYT 651
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 376 CMITDFPEPDITVEWQWNGOAPEN-YKNTQPIQMDT-DGSYFYVYSKLVNOKSWEAGNTFI 433

QY 652 CEATHK 657
PI ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 434 CSVLHE 439

RESULT 24
ID AAE18372 standard; Protein; 438 AA.
AC AAE18372;
XX AAE18372;
DT 07-MAY-2002 (first entry)
XX
DE Human penton base monoclonal antibody, DAV-1 heavy chain fragment.
XX
KW Human; bifunctional molecule; monoclonal antibody; gene therapy; cancer;
KW vascular disorder; diabetic retinopathy; restenosis; ophthalmic disorder;
KW hyperproliferative disorder; hormonal disorder; DAV-1 heavy chain;
KW cytostatic; vasotropic; ophthalmological.
XX
OS Homo sapiens.
XX
PN WO200204522-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-EP07878.
XX
PR 10-JUL-2000; 2000US-0613017.
XX
PA (NOVS ) NOVARTIS AG.
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
```

```
PA (SCRI ) SCRIPPS RES INST.
XX Nemerow GR, Li E;
DB WPI: 2002-171707/22.
DR N-PSDB; AAE18372.
XX
PT New bifunctional molecules comprising an antibody or its
PT antigen-binding portion, and a targeting agent, useful for e.g. gene
PT therapy, or for promoting Adenoviral vector-mediated gene delivery to
PT cells lacking av integrins
XX
PS Claim 10; Page 96; 106pp; English.
XX
CC The present invention relates to a bifunctional molecule comprising an
CC antibody or its antigen-binding portion, and a targeting agent where the
CC antibody specifically binds to an antigen in a protein that binds to av
CC integrin, and the targeting agent specifically binds to a cell surface
CC protein that activates the phosphatidylinositol 3 (PI3K) signalling
CC pathway. The bifunctional molecules are useful for gene therapy, for
CC promoting Adenoviral (Ad) vector-mediated gene delivery to cells lacking
CC av integrins, for enhancing Ad binding and internalisation, and in gene
CC delivery of by fibreless adenovirus particles. The bifunctional molecules
CC permit targeting of viral and bacterial vectors to cells that express
CC vascular disorders, diseases that can be targeted include cancers,
CC disorders, hyperproliferative disorders, restenosis, ophthalmic
CC disorders, hyperproliferative disorders, and hormonal disorders.
CC The present sequence is human penton base monoclonal antibody, DAV-1
CC heavy chain fragment.
XX
SQ Sequence 438 AA;
Query Match 29.0%; Score 1022.5; DB 23; Length 438;
Best Local Similarity 37.4%; Pred. No. 1.4e-53;
Matches 249; Conservative 50; Mismatches 111; Indels 255; Gaps 14;
QY 1 EVQLQQSGDPLVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGRINPNNGVTLY 60
DB 20 EVQLQQSGPELVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGYIYYPKGTGY 79
QY 61 NQFKDKATLTVDKSTTAYMELRLTSDSAVYVCARSTMTNVMYDYGQGTSTVSS 120
DB 80 NQFKSKATLTDTSSNTAYMELRLTSDSAVYVCARG-----IAYWGGTLTVSA 132
QY 121 AKTTPSVYPLAPGSAQAQNSWVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
DB 133 AKTTPSVYPLAPGSAQAQNSWVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 192
QY 181 LYTSSSVTPSPSTWPEVTCNVAHPASSTKVDDKIVPRDGGPSEKSEENKDLRKK 240
DB 193 LYTSSSVTPSPSTWPEVTCNVAHPASSTKVDDKIVPRDCG----- 235
QY 241 SELQGTALGNLKOIYYNYSKAITSEKSDAQFLTNTLLFKGFFTGHPWYNDLLVLDGSTA 300
DB 236 CKPCICTVEVSSVFIFPPK-----PKDVLITL----- 264
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTGPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKO 360
DB 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWEVD--- 289
QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVORGLIVF 411
DB 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHODWLNKKEKCRVNSAAPPAPIEK----- 343
QY 412 HSEGSTVSVDLFDAGQGYPDTLRLIYRDNTTISLSISLYLYTTSIVMTQTPTSLLV 471
DB 344 -----TISK-T----- 348
QY 472 SAGDRVTITCKASQSVNSDVAMVYQKPGSQKLLISYTSRYAGVPDRFSGSGYGTDFTL 531
DB 349 ----- 348
QY 532 TISSVQAEDAAYVFCQDYNSPPTFGGKTKLEIKRADAAPTYSIRPPSSQELTSGGASVY 591
DB 349 -----KGRPKAPQVITTPPPKQEMAKOKVSLT 375
QY 592 CFLNNFYPKIDINVKWKIDGSRQNGVLNSWTDQDSKDYSTMSSTLTLRKDEYERHNSYT 651
DB 376 CMITDFPEPDITVEWQWNGOAPEN-YKNTQPIQMDT-DGSYFYVYSKLVNOKSWEAGNTFI 433
QY 652 CEATHK 657
DB 434 CSVLHE 439
```

Qy	529	FTLTISVQAEADAAVYFCQDDYNSPPTFGGTTKLEIKRADAAPTVSIFPPSSQLTSGGA	588
Db	63	FTFTISTVQAEADLAAYFCQDDYSSPYTFGGGTTKLEIKRADAAPTVSIFPPSSQLTSGGA	122
Qy	589	SVVCFLNFPKDLINVKWKIDGSRQNGVLNSWTDQDSKDYTSMSSTLTLTCKDEYERHN	648
Db	123	SVVCFLNFPKDLINVKWKIDGSRQNGVLNSWTDQDSKDYTSMSSTLTLTCKDEYERHN	182
Qy	649	SYTCEATHKTSTSPVSKFNRRNE 671	
Db	183	SYTCEATHKTSTSPVSKFNRRNE 205	
RESULT 26			
AAR66758			
ID	AAR66758	standard; Protein: 465 AA.	
XX	AAR66758;		
XX	01-SEP-1995	(first entry)	
XX	Anti-tobacco mosaic virus monoclonal Ab heavy chain.		
XX	Tobacco mosaic virus; TMV; monoclonal antibody;		
XX	heavy chain; virus-resistant plants; biofarming.		
XX	Synthetic.		
XX	Key	Location/Qualifiers	
XX	Peptide	1..19	
XX	Peptide	/label= leader	
XX	Peptide	20..465	
XX	Peptide	/label= mat_peptide	
XX	Domain	20..128	
XX	Domain	/note= "variable heavy domain"	
XX	Domain	129..141	
XX	Domain	/note= "J heavy 4 domain"	
XX	Domain	142..465	
XX	Domain	/note= "constant heavy domain"	
XX	JP06319396-A.		
XX	22-NOV-1994.		
XX	07-MAY-1993;	93JP-0131208.	
XX	07-MAY-1993;	93JP-0131208.	
XX	(NISR) JAPAN TOBACCO INC.		
XX	(KURS) KURARAY CO LTD.		
XX	WPI: 1995-040220/06.		
XX	N-PSDB: AAQ79930.		
XX	Transformed plant producing animal-derived anti-virus antibody -		
XX	esp. tobacco plants producing anti-tobacco mosaic virus		
XX	monoclonal antibody		
XX	Example 2; Pages 14-15; 26pp; Japanese.		
XX	AAQ79929 and AAQ79930 encode AAR66757 and AAR66758, the light and heavy		
XX	chains of an animal derived anti-tobacco mosaic virus (TMV)		
XX	monoclonal antibody. The cDNAs were incorporated into a Ti		
XX	plasmid vector, which was incorporated into A. tumefaciens.		
XX	The resultant plant expression vector was used to transform		
XX	tobacco plants, making them TMV resistant, the plants could		
XX	also be biofarmed for the prodn. of anti-virus antibodies.		
XX	Sequence	465 AA;	
Query Match		28.9%; Score 1018; DB 16; Length 465;	
Best Local Similarity	36.3%;	Pred. No. 2.8e-53;	
Matches	244; Conservative	60; Mismatches 109; Indels 260;	Gaps

Qy 272 FLTN-----TLFKGFTGHPWYNLDLLVLDGTA-ATSEYEGSS--VDLYGAYGYQCAGG 324
 Db 378 MTKNOVSLTCLVKGYF-----SDIAVEWESNGOPENNYKTPPVLDSGSGFFLYSKL-- 430
 Qy 325 TPNKTACHYGGV---TLHD--NNRLTEEK-----KVPINLWIDGQTTVPIDKV 368
 Db 431 TVDKSRMOQGNVFCSCVMHEALHNNHYTKSLSPGKLGGSGVOLVQSGSELKPKPGASV 490
 Qy 369 RTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRGLIVFHSSEGSTVSYDLFDAQG 428
 Db 491 KISK-----ASCYTDDYGMNVKQAPCOGLKWMGWINTYTGESYVD--DFKG 538
 Qy 429 QYPTDLLRIYRNTTISTSLIS-----LYLYT----- 457
 Db 539 RF-----VFLSDTSVSAAYLQISSLKAEDTATYFCARRGFYAMDYWGQGTVTTVSSGGG 592
 Qy 458 -----TSIVMTQTPTSLLYSAGDRVITITCKASQSVNDVAVYQKPGQSPKLLI 506
 Db 593 GSGGGGGGGSDIVLTQSPATMSASGERVTLTCSASSSIY-MFWYHQRPQOSPRLLI 651
 Qy 507 SYTSRYSYAGVDRFSGSGYGTDFTLTITSSVOAEDAAYVFCQDYNSPPTFGGQTKLEIKR 566
 Db 652 YDTSNLASGVPAFSGSGSGSYSLTISRMEPEPATYFCHQSSSYPTFGGQTKLEIKR 711

RESULT 28
 AAY55081
 ID AAY55081 standard; Protein; 626 AA.
 XX
 AC AAY55081;
 DT 25-FEB-2000 (first entry)
 DE Single chain Fv protein sequence shPM1-kappa-BVGS3.
 KW Gene isolation; membrane-bound protein; fusion protein; drug production;
 KW antigen-binding cell; secreted functional protein; antigenic protein;
 KW protein isolation; diagnosis; ScFv.
 XX Synthetic.
 XX WO9960113-A1.
 XX 25-NOV-1999.
 XX 30-APR-1999; 99WO-JP02341.
 PR 20-MAY-1998; 98JP-0138652.
 XX 01-OCT-1998; 98JP-0279876.
 PA (CHUS) CHUGAI SEIVAKU KK.
 XX Tsuchiya M, Saito M, Ohtomo T;
 DR WPI: 2000-039382/03.
 XX N-PSDB: AAZ40316.
 PT Efficient and selective isolation of a gene encoding membrane protein
 PT with low or no antigenic binding activity, for diagnosis, study of, and
 PT production of drugs treating abnormal functions of the protein -
 XX
 FS Example 7; Page 103-109; 120pp; Japanese.
 XX This sequence represents a single chain Fv (ScFv) sequence.
 CC The invention relates to a method for isolating a gene encoding a
 CC membrane-bound protein, comprising introducing a vector into a cell,
 CC contacting an antigen with the cell expressing the fused protein encoded
 CC by the vector on its surface to select an antigen-binding cell, and
 CC isolating the cDNA. The vector contains DNA encoding a secreted
 CC functional protein with antigenicity and binding affinity, and a cDNA
 CC ligated to DNA downstream of the 3' end of the coding sequence. The
 CC method can be used to isolate a membrane-bound protein for diagnosis and
 CC study. It can also be used for producing drugs treating abnormal

CC functions of the protein. Such a technique is efficient and selective,
 CC which is different from the prior-art transmembrane trap (TMT) method
 CC wherein an epitope recognised by an antibody is carried in a fused
 CC protein.
 XX
 SQ Sequence 626 AA;
 Query Match 27.8%; Score 977.5; DB 21; Length 626;
 Best Local Similarity 35.2%; Pred. No. 1.le-50;
 Matches 253; Conservative 92; Mismatches 216; Indels 157; Gaps 22;
 Qy 1 EVQLQQSPDLVKPGASVKISCKASGYFTG-YMHVWYKSPGKLEWIGRINPNNGVTL 59
 Db 20 QVQLQESGPGLVPRPSQTLSTCTVSGYITSDHAWSVRQPPGRLGLEWIGVIS-YSGITT 78
 Qy 60 YNQRKFKDATLVDSKSTTAYMELRSLTSEDSAVYVCARSTMTINVMYDWCQGSVTVS 119
 Db 79 YNPSLKSRVTMLRDTSKNQFSLRSSVTAADTAVYCARSLARTT-AMDYWGQGSGLTVS 137
 Qy 120 SA-----KTPPSVYPLAPGSAOTNSMVTLGCL-----VKGYF----- 153
 Db 138 SGGGGGGGGGGSDIQMTQSPS-----SLSASVGDRTVITCRASODISSYLNWYQQ 191
 Qy 154 ---PEPVTVTWNSGSLSGVHTFPAVLQSDLYTLSSSVTVPSSTWSPSTVTT--CNAIHP- 207
 Db 192 KPGKAPKLLIYVTSRLHSGV---PSRFSGSGSGTDTFTTI-SSLQPEDTATYCCQGNLT 247
 Qy 208 ---ASSTKVDKIVPRDSGGSPSEKSEINEKDLRKKSELOCTALGNLQIYIYNSKAIT 263
 Db 248 PYTFGGQTKVEIKGGGGGGGGGGSGVD-----SQVQLQESGPGLV----- 291
 Qy 264 SSEKSADQFLTNTLLFKGF--FTGHPWYNLDLLVLDGTAATSEYEGSSVDLYGAYGYQC 321
 Db 292 ---PSQTLSTCTVSGYSITSDHAW-----SNVRQPPGRLGLE-----WIGY-- 329
 Qy 322 AGGPNKTAHYGGVTLHD---NNRLTEEKVPIINLWIDGQTTVPIDKVTSKKEVTVQ 378
 Db 330 -----ISYSGITTYNPSLKSRTVTLR-----DTSKKNQFSLR 360
 Qy 379 ELDLQAR---HYLHGKFGLYNSDSFGGKVORGLIVFHSSEGSTVSYDLFDAQGYQPTL 434
 Db 361 LSSVTAADTAVYCARSLARTTAMDYWG---OGSLVTVSSGGGGGGGGGG----- 411
 Qy 435 LRIYRDNNTTISSTLSLSILYLYTTSIVMTQTPTSLLLVSAGDRVITITCKASQSVNDVAVY 494
 Db 412 -----SDIQMTQSPSSLSASVGDRTVITCRASODISSYLNWY 448
 Qy 495 QOKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDFTLTISVVQAEADAAYFCQDYNPSP 554
 Db 449 QOKPGKAPKLLIYVTSRLHSGVPSRFSGSGSGTDTFTTISLQPEDTATYCCQGNLTLPY 508
 Qy 555 TFGGQTKLEIKRADAAPTVSIFFPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSEK 614
 Db 509 TFGGQTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQKVDNALQ 568
 Qy 615 NGVLNSWTDQDSKSTYSMSSTLTLTQDEYERHNSVTCETHKSTSPIVKSFNRNES 672
 Db 569 GNSQESVTEQDSKSTYSLSSTLTLSKADYKHKYACEVTHQGLSSPYTSFNRGES 626

RESULT 29
 AAY42294
 ID AAY42294 standard; Protein; 243 AA.
 XX
 AC AAY42294;
 XX
 DT 06-DEC-1999 (first entry)
 XX
 DE Anti-5T4 secreted single chain antibody Fv fragment.
 XX Cytochrome; targeting; localisation; cancer; tumour; prodrug; reduction;
 KW nucleus.


```
CC arthritis, or conditions associated with hypoxia, hypoglycemia or
CC ischemia, or to deliver antibiotics, antiviral agents, analgesics,
CC anesthetics, anti-inflammatories, antineoplastic agents and diagnostic
CC agents. LD optimize activity of PAD, e.g. by delivering it to selected
CC locations or by delivering it to neighboring cells (bystander effect),
CC and allow a reduction in dose of prodrug, and thus of systemic side-
CC effects. Nucleic acids encoding the agent may be expressed selectively
CC in hypoxic cells. The present sequence represents the single chain
CC variable antibody fragment against the tumor antigen 5T4 (5T4 scFv).
CC 5T4 scFv is used in the construction of a fusion protein comprising
CC 5T4 scFv and a human P450 reductase derivative alp450R.
XX
SQ Sequence 243 AA;
Query Match 27.6%; Score 972.5; DB 20; Length 243;
Best Local Similarity 40.1%; Pred. No. 6.9e-51;
Matches 227; Conservative 3; Mismatches 13; Indels 323; Gaps 5;
Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYMHVWVKQSPGKGLIEWIGRINPNNGVTLY 60
Db 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYMHVWVKQSHGKSLIEWIGRINPNNGVTLY 60
Qy 61 NQKFKDKATLTVDKSSTTAYMELRSLTSEDSAVYYCARSTMTITNYVMDYWGQGTSTVTVSS 120
Db 61 NQKFKDKAILTVDKSSTTAYMELRSLTSEDSAVYYCARSTMTITNYVMDYWGQVTSVTVS- 119
Qy 121 AKTTPPSVYPLAPGSAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
Db 120 ----- 119
Qy 181 LYTSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIVPRDSGGPSEKSEINEKDLRKK 240
Db 120 -----SGG----- 122
Qy 241 SELOQTALGNLKIYYNYSKAITSEKSAQDLTNTLLFKGFFTHPHYNDLLVDLGSTA 300
Db 123 ----- 122
Qy 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGCVTLHDNNRLTEKKVPINLWIDGKQ 360
Db 123 -----GGSGGGGT-----GG----- 132
Qy 361 TTVPIDKVKTSKKEVTVQELDLQARHYLHGKFLYNSDFGKVGORGLIVFHSSEGSTVS 420
Db 133 -----GG----- 134
Qy 421 YDLFDAQGOYPTDLLRIYRDNWTTISSTLSLSLYLTTSIVMTQTPTSLYSAGDRVTIT 480
Db 135 -----SSIVMTQTPTSLYSAGDRVTIT 157
Qy 481 CKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVPDRFSGSGYGTDTLTITSSVQAE 540
Db 158 CKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVPDRFSGSGYGTDTLTITSSVQAE 217
Qy 541 AAVYFCQQDYNSPPTFGGKTLEIKR 566
Db 218 LAVYFCQQDYNSPPTFGGKTLEIKR 243
RESULT 31
AAW86002
ID AAW86002 standard; Protein: 243 AA.
XX
AC AAW86002;
XX
DT 15-MAR-1999 (first entry)
DE Murine anti-5T4 antigen monoclonal antibody scFv.
XX
KW Tumour interacting protein: cancer; gene therapy: vector;
5T4 antigen; monoclonal antibody; single chain antibody; scFv;
mouse; 5T4scFv.1.
XX
```

```
OS Chimeric - Mus sp.
OS Chimeric - synthetic.
FH Key Location/Qualifiers
ET Misc-difference 169 /note= "encoded by GDT"
XX
PN W09855607-A2.
XX
PD 10-DEC-1998.
XX
PF 04-JUN-1998; 98WO-GB01627.
XX
PR 04-JUL-1997; 97GB-0014230.
PR 04-JUN-1997; 97GB-0011579.
PR 20-JUN-1997; 97GB-0013150.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Bebbington CR, Carroll MW, Ellard FM, Kingsman SM;
Myers KA;
XX
DR WPI; 1999-059910/05.
DR N-PSDB; AAW80290.
XX
PT New vector encoding a tumour interacting protein for treating cancer
PT - contains a desired nucleotide sequence and/or protein which
PT recognises tumours, and is used as a gene delivery system to treat
PT cancer
XX
PS Example 1; Fig 1A; 82pp; English.
XX
CC This amino acid sequence comprises an scFv, termed 5T4scFv.1,
CC comprising the heavy chain variable region (VH) from the murine 5T4
CC monoclonal antibody followed by a 15-amino acid flexible linker and
CC the light chain variable region (VL) of the mouse 5T4 antibody. The
CC trophoblast cell surface antigen defined by monoclonal antibody 5T4
CC is expressed at high levels on the cells of a wide variety of human
CC tumours. 5T4scFv.1 DNA (see AAW80290) can be used to construct
CC single-chain antibodies (see AAW86003) and scFv fusion constructs
CC (see AAW86004-05). The invention relates to a vector comprising a
CC nucleotide sequence coding for a tumour interacting protein (TIP)
CC and optionally a nucleotide sequence of interest (NOI) which
CC encodes a protein of interest (POI), the vector being capable of
CC delivering the NOI and/or POI to the tumour recognised by the TIP.
CC Delivery can be in vivo or ex vivo. The vector is used to treat
CC cancer, and may also used as a gene delivery system for introducing
CC at least 1 gene encoding a TIP (preferably a tumour binding protein)
CC into a haematopoietic cell lineage.
XX
SQ Sequence 243 AA;
Query Match 27.6%; Score 972.5; DB 20; Length 243;
Best Local Similarity 40.1%; Pred. No. 6.9e-51;
Matches 227; Conservative 3; Mismatches 13; Indels 323; Gaps 5;
```

```
Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYMHVWVKQSPGKGLIEWIGRINPNNGVTLY 60
Db 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYMHVWVKQSHGKSLIEWIGRINPNNGVTLY 60
Qy 61 NQKFKDKATLTVDKSSTTAYMELRSLTSEDSAVYYCARSTMTITNYVMDYWGQGTSTVTVSS 120
Db 61 NQKFKDKAILTVDKSSTTAYMELRSLTSEDSAVYYCARSTMTITNYVMDYWGQVTSVTVS- 119
Qy 121 AKTTPPSVYPLAPGSAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
Db 120 ----- 119
Qy 181 LYTSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIVPRDSGGPSEKSEINEKDLRKK 240
Db 120 -----SGG----- 122
Qy 241 SELOQTALGNLKIYYNYSKAITSEKSAQDLTNTLLFKGFFTHPHYNDLLVDLGSTA 300
```

Db 123 ----- 122
Qy 301 ATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
Db 123 -----GGSGGGGT----- 132
Qy 361 TTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVGQVGLIVFHSSEGSTVS 420
Db 133 -----GG----- 134
Qy 421 YDLFDAQGQYPTDLLRIYRDNTTSSLSISLYTTSIVMTQTPTSLVLSAGDRVTIT 480
Db 135 -----SSIVMTQTPTFLVLSAGDRVTIT 157
Qy 481 CKASQSVSNDAVWYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTSSVQAE 540
Db 158 CKASQSVSNDAVWYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTSSVQAE 217
Qy 541 AAVYFCQDYNSPPTFGGKLEIKR 566
Db 218 LAVYFCQDYNSPPTFGGKLEIKR 243
RESULT 32
AAB83835
ID AAB83835 standard; Protein; 243 AA.
AC AAB83835;
XX
XX
XX 23-JUL-2001 (first entry)
XX
XX Amino acid sequence of a 5T4 ScFv designated 5T4ScFv.1.
XX Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
XX hypersensitivity; autoimmune disease; central nervous system disorder;
XX Parkinson's disease; periodontal disease; cardiopulmonary disease;
XX cardiovascular disease; gastrointestinal disorder; infection; diabetes;
XX Helicobacter-related disease; immune disorder.
OS Synthetic.
OS Mus sp.
FH
FH Key Location/Qualifiers
FT Misc-difference 169
FT /note= "Ala encoded by GDT"
XX
XX WO200136486-A2.
XX
XX 25-MAY-2001.
XX
XX 13-NOV-2000; 2000WO-GB04317.
XX
XX 18-NOV-1999; 99WO-GB03859.
XX 15-FEB-2000; 2000GB-0003527.
XX 02-MAR-2000; 2000GB-0005071.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;
PI Myers KA;
XX
XX WPI: 2001-343805/36.
XX N-PSDB; AAF89729.
XX
XX Use of single chain antibody capable of recognizing a disease
XX associated molecule for manufacturing a medicament for preventing
XX and/or treating a disease condition associated with disease associated
XX molecule
XX
XX Claim 3; Fig 1; 118pp; English.
PS
XX The specification describes the use of a single chain antibody (ScFv),

CC which is capable of recognizing a disease associated molecule in the
CC manufacture of a medicament for the prevention and treatment of a
CC disease condition. The ScFv antibody is useful in the manufacture of a
CC medicament, for affecting a disease in vivo, for preparing a
CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
CC treatment of a disease. The ScFv antibody is also useful for
CC treating inflammatory diseases including arthritis, hypersensitivity,
CC autoimmune diseases, cancers, central nervous system disorders,
CC including Parkinson's disease, periodontal diseases, cardiopulmonary
CC diseases, cardiovascular diseases, gastrointestinal disorders,
CC infections, diabetes, Helicobacter-related diseases, and other immune
CC disorders. The present sequence represents a 5T4 ScFv of the invention.
CC The antibody comprises the VH and VL regions from murine 5T4 monoclonal
CC antibody, joined by a linker sequence.
XX
XX
XX Sequence 243 AA;
Query Match 27.6%; Score 972.5; DB 22; Length 243;
Best Local Similarity 40.1%; Pred. No. 6.9e-51;
Matches 227; Conservative 3; Mismatches 13; Indels 323; Gaps 5;
Qy 1 EVQLQQSGPDLVKGASVKISCKASGYSTFGYYMHVWKQSPCKGLEWIGRINPNNGVTLY 60
Db 1 EVQLQQSGPDLVKGASVKISCKASGYSTFGYYMHVWKQSPCKGLEWIGRINPNNGVTLY 60
Qy 61 NQKFDKATLTVDKSSSTAYMELRSLTSEDSAVYYCARSTMTNVMYDYGQGTSTVYSS 120
Db 61 NQKFDKATLTVDKSSSTAYMELRSLTSEDSAVYYCARSTMTNVMYDYGQGTSTVYSS 119
Qy 121 AKTTPPSYIPLAGSAAQTNSMVTGLCLVKGYPFEPVTVTNWSGLSSGVHTFFAVLQSD 180
Db 120 ----- 119
Qy 181 LYTSSSVTPSSWPSETVTCNVHPASSTKVDKIVPRDSGGPSEKSEINEKDLRKK 240
Db 120 -----SGG----- 122
Qy 241 SELOCTALGNLQIYYNNSKAITSEKSADQFLTNTLLFKGFTGHPWYNDLLVDLGSTA 300
Db 123 ----- 122
Qy 301 ATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
Db 123 -----GGSGGGGT----- 132
Qy 361 TTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVGQVGLIVFHSSEGSTVS 420
Db 133 -----GG----- 134
Qy 421 YDLFDAQGQYPTDLLRIYRDNTTSSLSISLYTTSIVMTQTPTSLVLSAGDRVTIT 480
Db 135 -----SSIVMTQTPTFLVLSAGDRVTIT 157
Qy 481 CKASQSVSNDAVWYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTSSVQAE 540
Db 158 CKASQSVSNDAVWYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTSSVQAE 217
Qy 541 AAVYFCQDYNSPPTFGGKLEIKR 566
Db 218 LAVYFCQDYNSPPTFGGKLEIKR 243
RESULT 33
AAY44176
ID AAY44176 standard; Protein; 214 AA.
XX
XX AAY44176;
XX
XX 01-FEB-2000 (first entry)
XX
XX MAb Fab13p5 heavy chain protein sequence.
XX
XX Peptide ligand; affinity; p24; human immune deficiency virus-1; HIV-1;
KW

KW light chain; heavy chain; Fab; monoclonal antibody; hypervariable region;
 KW infection.
 XX
 OS Mus sp.
 OS FR2777285-A1.
 PN
 XX
 PD 15-OCT-1999.
 XX
 XX 10-APR-1998; 98FR-0004876.
 XX
 XX 10-APR-1998; 98FR-0004876.
 XX
 XX (INMR) BIO MERIEUX.
 XX
 XX Novelli RA, Monaco S, Piga N, Berthet C, Mallet F, Cusack S;
 PI Chassaigne V;
 XX
 XX WPI: 1999-593428/51.
 DR N-PSDB; AA228805.
 XX
 PT New peptide ligand specific for p24 of human immune deficiency virus
 PT contains hypervariable regions of antibody 13B5, used for diagnosing
 PT HIV infection
 XX
 PS Claim 2: Page 20-21; 27pp; French.
 XX
 CC The invention relates to a peptide ligand with specific affinity for
 CC the p24 protein of human immune deficiency virus-1 (HIV-1) comprising
 CC at least one peptide strand corresponding to the N-terminal region of
 CC the light and/or heavy chain of the Fab fragment of monoclonal antibody
 CC 13B5 in which: (i) the light chain includes three hypervariable regions
 CC (HVR) at amino acid (aa) positions 24-33, 49-55 and 88-95 of AAY44175;
 CC and (ii) the heavy chain includes three HVR at aa positions 26-35,
 CC 49-65 and 99-109 of this sequence. The peptide ligands are reagents
 CC for detecting p24 (by standard immunoassays) in biological samples,
 CC specifically for diagnosis of HIV-1 infection or can be used to treat
 CC HIV-1 infections.
 XX
 SQ Sequence 214 AA:
 Query Match 27.5%; Score 970; DB 20; Length 214;
 Best Local Similarity 85.0%; Pred. No. 8.3e-51;
 Matches 182; Conservative 15; Mismatches 17; Indels 0; Gaps 0;
 QY 1 EVLOQSGDPLVKPGASVKISCKASYSTGYMHVWVKSPGKLGIEWIGRIINPNNGVTLY 60
 DB 1 EVLOQSGAELARPGASVKMSCKASYTFTSYMHVVKRPGOGLEWIGYINPSSGYSNY 60
 QY 61 NQKFKDKATLTVDKSSSTTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGQGTSTVTVSS 120
 DB 61 NQKFKDKATLTADKSSSTAYMQLSLSLTSDSAVYYCSRVPVRLGYNFDYWGQGTSLTVSS 120
 QY 121 AKTTPPSVYPLAPGSAQAQNSVMTLGCLVKGYPPEPTVTWNSSGSLSSGVHTFPVQLQSD 180
 DB 121 AKTTPSVYPLAPGSAQAQNSVMTLGCLVKGYPPEPTVTWNSSGSLSSGVHTFPVQLQSD 180
 QY 181 LYTSSSVTPSPSTWSPSETVTCNVAHPASSTKVD 214
 DB 181 LYTSSSVTPSPSTWSPSETVTCNVAHPASSTKVD 214
 RESULT 34
 ID AA086004
 XX AA086004 standard; Protein; 488 AA.
 AC AA086004;
 XX
 XX 15-MAR-1999 (first entry)
 XX Human B7-1.5T4.1 protein fusion, specific for human 5T4.
 XX Tumour interacting protein; cancer; gene therapy; vector;
 KW

KW 5T4 antigen; monoclonal antibody; single chain antibody;
 KW mouse; human; B7-1; co-stimulatory molecule.
 XX
 OS Chimeric - Mus sp.
 OS Chimeric - Homo sapiens.
 OS Chimeric - synthetic..
 XX
 PN WO9855607-A2.
 XX
 XX 10-DEC-1998.
 XX
 XX 04-JUN-1998; 98WO-GB01627.
 PF
 XX 04-JUL-1997; 97GB-0014230.
 PR 04-JUN-1997; 97GB-0011579.
 PR 20-JUN-1997; 97GB-0013150.
 XX
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 XX Bebbington CR, Carroll MW, Ellard FM, Kingsman SM;
 PI Myers KA;
 XX
 XX WPI: 1999-059910/05.
 DR N-PSDB; AAV80292.
 XX
 PT New vector encoding a tumour interacting protein for treating cancer
 PT - contains a desired nucleotide sequence and/or protein which
 PT recognises tumours, and is used as a gene delivery system to treat
 PT cancer
 XX
 XX Example 5; Fig 2; 82pp; English.
 PS
 CC This is the amino acid sequence of B7-1.5T4.1, a fusion protein
 CC comprising the extracellular domain (amino acids 1-215) of human
 CC co-stimulatory molecule B7-1 joined via a flexible peptide linker
 CC to an scFv (see AAW86002) derived from murine 5T4 monoclonal
 CC antibody. B7-1.5T4.1 cDNA (see AAV80292) can be inserted into vector
 CC pCI to allow expression of the fusion protein in mammalian cells.
 CC The trophoblast cell surface antigen defined by 5T4 is expressed at
 CC high levels on the cells of a wide variety of human tumours. The
 CC invention relates to a vector comprising a nucleotide sequence
 CC coding for a tumour interacting protein (TIP) and optionally a
 CC nucleotide sequence of interest (NOI) which encodes a protein of
 CC interest (POI), the vector being capable of delivering the NOI
 CC and/or POI to the tumour recognised by the TIP. Delivery can be in
 CC vivo or ex vivo. The vector is used to treat cancer, and may also
 CC used as a gene delivery system for introducing at least 1 gene
 CC encoding a TIP (preferably a tumour binding protein) into a
 CC haematopoietic cell lineage. B7-1 is expected to bind specifically
 CC to CD28 and CTLA-4 present on human T-cells.
 XX
 SQ Sequence 488 AA;

Query Match 27.5%; Score 967.5; DB 20; Length 488;
 Best Local Similarity 40.0%; Pred. No. 3.2e-50;
 Matches 226; Conservative 3; Mismatches 13; Indels 323; Gaps 5;
 QY 1 EVLOQSGDPLVKPGASVKISCKASYSTGYMHVWVKSPGKLGIEWIGRIINPNNGVTLY 60
 DB 247 EVLOQSGDPLVKPGASVKISCKASYSTGYMHVWVKSHGKSLIEWIGRIINPNNGVTLY 306
 QY 61 NQKFKDKATLTVDKSSSTTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGQGTSTVTVSS 120
 DB 307 NQKFKDKATLTVDKSSSTTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGQGTSTVTVSS- 365
 QY 121 AKTTPPSVYPLAPGSAQAQNSVMTLGCLVKGYPPEPTVTWNSSGSLSSGVHTFPVQLQSD 180
 DB 366 -----SGG----- 365
 QY 181 LYTSSSVTPSPSTWSPSETVTCNVAHPASSTKVDKIVPRDGGPSEKSEINEKDLRKK 240
 DB 366 -----SGG----- 368

QY 241 SELQGTALGNLQIYYNSKAITSEKSADQFLNTLLFKGFFTHPWNNDLLVLDGSTA 300
 Db ----- 368
 QY 301 ATSEYEGSSVDLYGAYGYQACGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
 Db ----- 378
 QY 361 TTVPIDKVKTSKKEVTQVQLDQARHYLHGKFGLYNSDSFGKVGQGLVLFHSSGEGSTVS 420
 Db ----- 380
 QY 421 YDLFDAQGOYPTDLLRIYRDNTTISSTLSLSLYLTTISIVMTQTPTSLLSVAGDRVTIT 480
 Db ----- 403
 QY 481 CRASQSVSNDVAMVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTISSVQAE 540
 Db ----- 463
 QY 541 AAVYFCQDDYNSPPTFGGCTKLEIK 565
 Db ----- 488
 QY 464 LAVYFCQDDYNSPPTFGGCTKLEIK 488
 Db -----
 RESULT 35
 AAB83836
 ID AAB83836 standard; Protein; 488 AA.
 AC AAB83836;
 XX
 DT 23-JUL-2001 (first entry)
 DE Amino acid sequence of a B7-1.5T4.1 fusion protein.
 XX
 KW Single chain antibody; ScFv; Inflammatory disease; arthritis; cancer;
 KW hypersensitivity; autoimmune disease; central nervous system disorder;
 KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
 KW cardiovascular disease; gastrointestinal disease; infection; diabetes;
 KW Helicobacter-related disease; immune disorder.
 XX
 OS Synthetic.
 OS Mus sp.
 OS Homo sapiens.
 XX
 PN WO200136486-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 13-NOV-2000; 2000WO-CB04317.
 XX
 PR 18-NOV-1999; 99WO-CB03859.
 PR 15-FEB-2000; 2000GB-0003527.
 PR 02-MAR-2000; 2000GB-0005071.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard PM;
 PI Myers KA;
 XX
 DR WPI; 2001-343805/36.
 DR N-PSDB; AAF89730.
 XX
 PT Use of single chain antibody capable of recognizing a disease
 PT associated molecule for manufacturing a medicament for preventing
 PT and/or treating a disease condition associated with disease associated
 PT molecule -
 XX
 PS Claim 3; Fig 2; 118pp; English.
 XX
 CC The specification describes the use of a single chain antibody (ScFv),
 CC which is capable of recognizing a disease associated molecule in the
 CC manufacture of a medicament for the prevention and treatment of a

CC disease condition. The ScFv antibody is useful in the manufacture of
 CC a medicament, for affecting a disease in vivo, for preparing a
 CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
 CC treatment of a disease. The ScFv antibody is also useful for
 CC treating inflammatory diseases including arthritis, hypersensitivity,
 CC autoimmune diseases, cancers, central nervous system disorders
 CC including Parkinson's disease, periodontal diseases, cardiopulmonary
 CC diseases, cardiovascular diseases, gastrointestinal disorders,
 CC infections, diabetes, Helicobacter-related diseases, and other, immune
 CC disorders. The present sequence represents a B7-1.5T4.1 fusion protein.
 CC This comprises the N-terminus of the 5T4 ScFv is fused after amino acid
 CC 215 of human B7-1.
 XX
 SQ Sequence 488 AA;
 Query Match 27.5%; Score 967.5; DB 22; Length 488;
 Best Local Similarity 40.0%; Pred. No. 3.2e-50;
 Matches 226; Conservative 3; Mismatches 13; Indels 323; Gaps 5;
 QY 1 EVQLQSGDPLVKPGASVKISCKASGYSTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60
 Db ||||| 323
 QY 247 EVQLQSGDPLVKPGASVKISCKASGYSTGYIMHWKQSHGKSLWIGRINPNNGVTLY 306
 Db ||||| 306
 QY 61 NQKFKDKATFLVDSSTTAYMELSLTSDSAVYYCARSTMITNYMDYWGQCTSVTVSS 120
 Db ||||| 365
 QY 307 NQKFKDKATFLVDSSTTAYMELSLTSDSAVYYCARSTMITNYMDYWGQCTSVTVSS- 365
 Db ||||| 365
 QY 121 AKTTPPSVYPLAPGAAQNTNSWTLGCLVKGYFPEPVTVWNSGSLSSGVHTFPAVLQSD 180
 Db ----- 365
 QY 181 LYTLSSTVTPSPSTWPEVTTCNVAHPASSTKVKKIVPRDSGGPSEKSEEINEKDLRKK 240
 Db ||||| 368
 QY 241 SELQGTALGNLQIYYNSKAITSEKSADQFLNTLLFKGFFTHPWNNDLLVLDGSTA 300
 Db ----- 368
 QY 301 ATSEYEGSSVDLYGAYGYQACGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
 Db ----- 378
 QY 361 TTVPIDKVKTSKKEVTQVQLDQARHYLHGKFGLYNSDSFGKVGQGLVLFHSSGEGSTVS 420
 Db ----- 380
 QY 421 YDLFDAQGOYPTDLLRIYRDNTTISSTLSLSLYLTTISIVMTQTPTSLLSVAGDRVTIT 480
 Db ----- 403
 QY 481 CRASQSVSNDVAMVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTISSVQAE 540
 Db ----- 463
 QY 541 AAVYFCQDDYNSPPTFGGCTKLEIK 565
 Db ----- 488
 QY 464 LAVYFCQDDYNSPPTFGGCTKLEIK 488
 Db -----
 RESULT 36
 AARL3061
 ID AARL3061 standard; Protein; 468 AA.
 XX
 AC AARL3061;
 XX
 DT 03-OCT-1991 (first entry)
 XX
 DE Monoclonal antibody OK3T heavy chain.
 XX
 KW OK3T; light chain; humanised antibodies; CDR-grafting.
 XX
 OS Mus musculus.

CC proteins containing two different Ab-derived units (i.e. to produce
CC bispecific antibodies). When a toxic protein is used in place of
CC phoA, the hybrid molecules can be used as cell-targeting
CC therapeutic agents.
XX
SQ Sequence 225 AA;

Query Match 27.3%; Score 961.5; DB 14; Length 225;
Best Local Similarity 88.6%; Pred. No. 2.9e-50;
Matches 194; Conservative 4; Mismatches 8; Indels 13; Gaps 5;

QY 459 SIWMTQTPSLVLSAGDRVTITCKASQSVNSDVAVYQKPGSPKLLIS--YTSRYAGV 516
||||| ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 13 SIWMTQTPKFLLSAGDRVTITCKASQSVNSDVAVW--KPGQSPKLLIQYVASSRYTCV 70

QY 517 PDRESGSGYGTDFTLTISVQAEADRAVYFCQDYNPPTFGGKLEIKRAAAPTVSIF 576
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 71 PDRETSQSGYGTDFTLTISVQAEADLAVYFCQDYS--YTFGGGKLEIKRAAAPTVSIF 129

QY 577 PPSSEQLTSGGASVVCFLNFPKIDNVKWKIDGSRQGVLSNWTDDQSKDSTYSMSST 636
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 130 PPSSEQLTSGGASVVCFLNFPKIDNVKWKIDGSRQGVLSNWTDDQSKDSTYSMSST 189

QY 637 LTLTKDEYERH----NSYTCEATHKTSTSPIVKSFNRRNE 671
||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 LTL----YERHTKDENSYTCEATHKTSTSPIVKSFNRRNE 224

RESULT 38
AAW15932
ID AAW15932 standard; Protein; 212 AA.
XX
AC AAW15932;
XX
DT 09-DEC-1997 (first entry)
XX
DE Antibody 7G12 heavy chain variable region.
XX
KW catalytic antibody; enantioselective hydrolysis; hybridoma;
KW ZAA7G12; ZAA3G2; ds.
XX
OS Mus.
XX
PN WO9629426-A1.
XX
PD 26-SEP-1996.
XX
PF 17-MAR-1995; 95WO-JP00462.
XX
PR 17-MAR-1995; 95WO-JP00462.
XX
PA (PROT-) PROTEIN ENG RES INST.
XX
PI Fujii I, Kinoshita K, Tanaka F;
XX
WPI; 1996-443199/44.
DR N-PSDB; AAT87815.
DR
XX
PT Catalytic antibody for enantioselective hydrolysis of amino acid
PT esters - also new hybridoma secreting the antibody
XX
PS Disclosure; Page 31-32; 49pp; Japanese.
XX
CC The patent discloses new catalytic antibodies which hydrolyse amino
CC acid esters enantioselectively. Preferably the esters are 4-nitro-
CC benzyl esters and the esterified amino acids are amino-protected.
CC Also disclosed are new hybridomas expressing the catalytic antibodies,
CC especially ZAA7G12 (FERM BP-4947) and ZAA3G2 (FERM BP-4946). The
CC antibodies are raised in mice using the compound p-nitrobenzyl
CC [4-carboxy-1-(benzyloxy)carbonylamino]-butyllphosphonate as haptens.
CC They are used for efficient resolution of racemic amino acids with
CC high optical selectivity, giving optically active amino acids useful
CC for the production of optically active drugs and chiral separation

CC agents.
CC The present sequence represents the variable region of the heavy chain
CC of antibody 7G12, produced by the hybridoma ZAA7G12.
XX
SQ Sequence 212 AA;

Query Match 27.3%; Score 960; DB 17; Length 212;
Best Local Similarity 84.7%; Pred. No. 3.3e-50;
Matches 183; Conservative 11; Mismatches 14; Indels 8; Gaps 1;

QY 6 QSGDPLVPGASVKISCKASGYSTFGYMHVWKQSPGKLEIGRINPNNGVTLYNQKEK 65
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2 ESGTELAKPGASVKRMSKASGYSTFTSYIHVWKVRPGQGLEWIGYINPSTDYTIQKEK 61

QY 66 DKATITVDKSSITAYMELRSLTSEDSAVYVCARSTMITNYVMDYWGQGTSTVSSAKTTP 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 62 DKATITADKSSSTAYMQLSSLTSEDSAVYVCV-----MKDYWGQGTSTVSSAKTTP 113

QY 126 PSVYPLAPGSAQAOTNSMTLGLVKGYFPEPVTVTWNSGSLSSGVHTFPVAVLQSDLYTLS 185
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 114 PSVYPLAPGSAQAOTNSMTLGLVKGYFPEPVTVTWNSGSLSSGVHTFPVAVLQSDLYTLS 173

QY 186 SSVTPSPSTWPSSETVTCNVVHPASSTKVKIVPRD 221
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 174 SSVTPSPSTWPSSETVTCNVVHPASSTKVKIVPRD 209

RESULT 39
AARI3203
ID AARI3203 standard; protein; 233 AA.
XX
AC AARI3203;
XX
DT 15-OCT-1991 (first entry)
XX
DE Staphylococcal enterotoxin A.
XX
KW SEA; cancer treatment; pyrogen; tumouricide.
XX
OS Staphylococcus aureus.
XX
PN WO9110680-A.
XX
PD 25-JUL-1991.
XX
PF 17-JAN-1991; 91WO-US00342.
XX
PR 17-JAN-1990; 90US-0466577.
XX
PA (TERM/) Terman D S.
XX
PI Terman DS;
XX
WPI; 1991-237984/32.
XX
PT Treating cancer with enterotoxin from Staphylococcus aureus -
PT administered by IV injection, having same tumoricidal activity
XX as Staphylococcal protein A without potential toxic reactions
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC SEA was isolated and purified from S. aureus. It can be used for
CC treating cancer, activating cytokine mediators and procoagulant
CC systems, augmenting natural killer cell cytotoxicity, etc. The
CC enterotoxin can be administered intravenously, optionally with
CC ibuprofen to attenuate toxic reaction to SEA. Synthetic
CC polypeptides having structural homology to SEA. Synthetic
CC are claimed, provided the homology includes statistically
CC significant sequence homology, alignment of Cysteine residues and
CC similar hydropathy profiles.
CC See also AARI3204-R13211.
XX
SQ Sequence 233 AA;

CC into them to form a conjugate. V-min is a ligand-binding variable
CC domain which comprises a framework region of a cyclically permuted
CC central beta-barrel, outer beta-sheet segments, and linker segments,
CC and complementarity determining regions. V-min, when conjugated, are
CC useful in radioimaging, as immunotoxins and in antibody directed
CC enzyme prodrug therapy, ie. for diagnosis or treatment of tumours.

XX
SQ Sequence 220 AA:

Query Match 27.1%; Score 953; DB 15; Length 220;

Best Local Similarity 84.4%; Pred. No. 9e-50;

Matches 184; Conservative 14; Mismatches 14; Indels 6; Gaps 1;

QY 460 IYVOTPTSLVLSAGDRTVITTCASQVSNLND-----VAWYQKQKPGQSPKLLISYTSRRY 513

DB 2 IYVOTGPSLSVLSAGERTVMSCKSSQSLNSGNKFNLAWYQKQKPGQSPKLLIYGASTRE 61

QY 514 AGVPRDFSGSGYGTDTLTITSSVQAEADAAYFCQDYNPPTFGGKLEIKRADAAPT 573

DB 62 SGVPRDFSGSGYGTDTLTITSSVQAEADLAAYVCQNDHSYPLTFGAGTLEIKRADAAPT 121

QY 574 SIFFPSSEQLTSGGASVVCFLNFPKIDNVKWKIDGSRQNGVLSNWTQDQSKDSTYSM 533

DB 122 SIFFPSSEQLTSGGASVVCFLNFPKIDNVKWKIDGSRQNGVLSNWTQDQSKDSTYSM 181

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

QY 634 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 671

DB 182 SSTLTITKDEYERHNSYTCETHKSTSPIVKSFNRNE 219

XX New immunoglobulin hybrid proteins - with immunoglobulin
PT fragments linked to dimeric protein, for diagnostic or
PT therapeutic use
XX
PS Example 1; Fig 3A; 37pp; French.

XX
CC A fragment of the heavy chain (VH + CH1) from the anti-snake small

CC neurotoxin monoclonal antibody M(alpha)2-3 was PCR-amplified from
CC hybridoma-derived cDNA using primers AAQ48039 and AAQ48040. A light
CC chain fragment (VL + CL) was amplified from the same source using
CC primers AAQ48041 and AAQ48042. The two amplified fragments were
CC inserted into the same vector; the H-chain fragment was inserted
CC (in-frame) between codons 6-7 of the phoA coding sequence and the
CC L-chain fragment was inserted into a cassette which contained a
CC phoA 5'-D sequence, a signal peptide and the first 6 codons of phoA.
CC The cassette was positioned between the termination codon and
CC the transcription termination sequence of phoA. The fusion
CC construct is expected to encode a hybrid protein comprising two
CC identical Ab-derived units. The invention also covers hybrid
CC proteins containing two different Ab-derived units (i.e. to produce
CC bispecific antibodies). When a toxic protein is used in place of
CC phoA, the hybrid molecules can be used as cell-targeting
CC therapeutic agents.

XX
SQ Sequence 459 AA:

Query Match 27.0%; Score 951.5; DB 14; Length 459;

Best Local Similarity 34.2%; Pred. No. 2.8e-49;

Matches 236; Conservative 62; Mismatches 126; Indels 267; Gaps 13;

QY 1 EVLOQSGPDLVKPGASVKISCKASGYSTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60

DB 20 QIQLOQSGPELVKPGASVKISCKASGYSTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 79

QY 61 NQKFKDKATLVKDSSTAYMELRLSLTSDSAVYYCARSTMTINTYMDYWGOGTSVTYSS 120

DB 80 NENFKGKATLVDTSSSTAYMQLSLTSEDATVYFCARAGMATATLLDYWGOGTTLTVSS 139

QY 121 AKTTPSVYPLAPGSAQAQNSMTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSD 180

DB 140 AKTTAPSVYPLAPVCGDVTGSSVTLGLVKGYFPEPVTLTWNSGSLSSGVHFFPAVLQSD 199

QY 181 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVDKKIYPRDGGSPGSEKSEINEKDLRKK 240

DB 200 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVDKKIYPRDGGSPGSEKSEINEKDLRKK 247

QY 241 SELOCTALGNLKOIYYNYSKAITSEKSAQDLTNTLLFKGFTHGHPWYNDLLDLGSTA 300

DB 248 ----- 247

QY 301 ATSEYEGSSVDLYGAYGYQCAGTGNKTCAGYGVTLHDNNRLTEKKVPINLWIDGKQ 360

DB 248 ----- 263

QY 361 TTVPIDKVKTSKKEVTQVQLDQARHLYLHGKFLYNSDSFGKGVQGLIVFHSSGTSVS 420

DB 264 ----- 282

QY 421 YDLFDAQGGYPTLLRIYRDNTTISSTLSISLYLYTTISIVMTQPTTLLVSAGDRVTIT 480

DB 283 CVVVDVSEDDPD-----VQISWFVNNVEVHTAQOT----- 313

QY 481 CKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVPRDFSGSGYGTDTLTISVQAEQ 540

DB 314 ----- 325

QY 541 AAVYFCQDYNPPTFGGKLEIKRAD-----AAPTYSIFPPSSEQLT 584

DB 326 SALPTQHQDWSGKEF-----KCKVNNKDLPAPIERTISKPGSVRAPQVYVLPPEEWT 381

QY 585 SGGASVVCFLNFPKIDNVKWKIDGSRQN-----GVLNSWTDQDQSKDSTYSMTLTL 639

DB ----- 639

Db 182 CLIVFHTSTEPSVNYDLFGAOGQVSNLTLRIYRDNKTINSENMHIDIYLYTS 233

RESULT 47

AAR45011 ID AAR45011 standard; protein; 233 AA.

XX AC AAR45011;

XX DT 08-JUN-1994 (first entry)

XX DE Staphylococcal enterotoxin SEA.

XX KW Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;

XX KW autoimmune disease; toxicity; Protein A; perfusion system.

XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers

XX FT Misc-difference 49

XX ET /note= "Given in the specification as O, no further details given"

XX PN WO9324136-A.

XX PD 09-DEC-1993.

XX PF 01-JUN-1993; 93WO-US05213.

XX PR 01-JUN-1992; 92US-0891718.

XX PA (STON/) STONE J L.

XX PA (TERM/) TERMAN D S.

XX XX Stone JL, Terman DS;

XX DR WPI; 1993-405418/50.

XX PT Use of staphylococcal enterotoxin(s) and homologues - for treating cancer in a patient or for the treatment of auto-immune diseases

XX PS Disclosure; Fig 1; 90pp; English.

XX CC The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs) which may be used in the methods of the invention for treating cancer in a patient. These SEs, and homologues of them, can be used as tumouricidal agents for treating cancers and autoimmune disease. They exhibit tumouricidal activity and toxicity identical to that observed for the Protein A perfusion system. They may be administered by i.v. injection.

XX SQ Sequence 233 AA;

Query Match 26.7%; Score 942; DB 14; Length 233;

Best Local Similarity 76.0%; Pred. No. 4.4e-49;

Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

Qy 236 SEKSEINEKDLRKKSELOGTALGNLKIYYNNKAITSSSEKSDQFLTNTLLFKGFTG 285

Db 1 SEKSEINEKDLRKKSELOGTALGNLKIYYNNKAITSSSEKSDQFLTNTLLFKGFTG 60

Qy 286 HPWYNDLLVGLGTAATSEYESSVDLYGAYGYOCAGGTGPNKTACMYGGVTLHDNNBLT 345

Db 61 HSWYNDLLVDFSDIDVYKGGKVDLYGAYGYOCAGGTGPNKTACMYGGVTLHDNNBLT 120

Qy 346 BEKKVPINLWDGKQTTVPIDKVKTSKREVTAVQELDLQARHYLHGKFLYNSDSFGGKVQ 405

Db 121 BEKKVPINLWDGKQTTVPLETYKTNKNVYQELDPQARYLQEKYLYNSDVPDGKVQ 180

Qy 406 RGLIVFHSSEGSTVSVDLFGAOGQVDPDTLLRIYRDNKTINSENMHIDIYLYTS 458

Db 181 RGLIVFHTSTEPSVNYDLFGAOGQVSNLTLRIYRDNKTINSENMHIDIYLYTS 233

RESULT 48

AAR76086

ID AAR76086 standard; Peptide; 219 AA.

XX AC AAR76086;

XX DT 21-NOV-1995 (first entry)

XX DE Mab 55.1 light chain.

XX KW Antigen binding structure; complementarity determining region; CDR; CA55.1; colorectal cancer; tumor-associated antigen; hybridoma; monoclonal antibody; Mab; immunotherapy; therapy; diagnosis; transgenic animal; transgenic plant; antibody engineering; humanized antibody; immunotoxin.

XX OS Mus sp.

XX PN WO9515382-A.

XX PD 08-JUN-1995.

XX PF 29-NOV-1994; 94WO-GB02610.

XX PR 03-JUN-1994; 94GB-0011089.

XX PR 03-DEC-1993; 93GB-0024819.

XX PA (ZENE) ZENECA LTD.

XX PI Blakey DC, Boot C, Copley CG, Hall SM, Paterson DS;

XX PI Rose MS, Wright AF;

XX DR WPI; 1995-215262/28.

XX PT Antigen binding structures containing CDRs recognising the CA55.1 antigen - produced by hybridomas and host cells, for use in the diagnosis and therapy of cancer

XX PS Claim 3; Page 98; 121pp; English.

XX CC An antigen binding structure is based on the CDRs (given in AAR76078-84) of the heavy (AAR76085) and light (AAR76086) chains of Mab 55.1 (ECACC 93081901), which recognises the colorectal tumor-associated antigen CA55.1. It is optionally humanized and in the form F(ab')₂, F(ab)', Fab, Fv, scFv or V-min, and is produced in transgenic animals or plants.

XX SQ Sequence 219 AA;

Query Match 26.7%; Score 941.5; DB 16; Length 219;

Best Local Similarity 83.0%; Pred. No. 4.4e-49;

Matches 181; Conservative 19; Mismatches 11; Indels 7; Gaps 2;

Qy 460 IVMTOTPTSLVSGADRYTITCKASQSVND-----VAVYQOKPGQSPKLLISYTSRY 513

Db 2 IVMSQSPSLAVSAGEKVTMCKSSQSLNSTRKNYLAWYQQRPQSPKLLIYWASTRT 61

Qy 514 AGVPDRFSGSGYGTDTLTITSSVQAEDAAVYFCQODYNSPPTFGGCTKLEIKRADAAPT 573

Db 62 SGVPDRFSGSGYGTDTLTITSSVQAEDLAIYCKQSY-TLRTFGGCTKLEIKRADAAPT 120

Qy 574 SIFPPSSQLTSGGASVVCFLNNFYPKDNVKKWKIDGSEGRQNGVLNSWTDDQSKDSTYS 633

Db 121 SIFPPSSQLTSGGASVVCFLNNFYPKDNVKKWKIDGSEGRQNGVLNSWTDDQSKDSTYS 180

Qy 634 SSTLTTLTKDEYERHNSYTCETHTKTSTSPIVKSFNRNE 671

Db 181 SSTLTTLTKDEYERHNSYTCETHTKTSTSPIVKSFNRNE 218

RESULT 49
AAW35373
ID AAW35373 standard; peptide; 233 AA.
XX
AC AAW35373;
XX
DT 20-APR-1998 (first entry)
XX
XX Staphylococcus enterotoxin SEA wild-type superantigen.
XX
XX SEE; SEA; Staphylococcus enterotoxin; superantigen; conjugate;
KW treatment; cancer; infection; autoimmune disease; antibody.
XX
XX Staphylococcus sp.
OS
OS W09735932-A1.
PN
XX 09-OCT-1997.
PD
XX 26-MAR-1997; 97WO-SE00537.
PF
XX 12-AUG-1996; 96US-0695692.
PR
PR 29-MAR-1996; 96SE-0001245.
XX
XX (PHAA) PHARMACIA & UPJOHN AB.
FA
XX Abrahmsen L, Antonsson P, Bjoerk P, Dohlisten M;
PI Forsberg G, Hansson J, Kalland T;
PI
XX WPI; 1997-503052/46.
DR
XX Conjugate of target seeking moiety and modified superantigen -
PT useful for activating the immune system to treat cancer, viral
PT infections, parasitic infestations and autoimmune diseases
XX
PS Claim 8; Pages 36-37; 58pp; English.

XX This is the wild-type Staphylococcus enterotoxin SEA superantigen. This
XX SEA superantigen can be modified to be used in a novel conjugate. The
CC novel conjugate comprises a target seeking moiety and a modified wild
CC type superantigen. The modified superantigen retains its ability to
CC activate a subset of T cells, even though 1 or more wild-type amino acid
CC residues in at least 1 region which functions in determining binding to
CC T cell receptor (TCR) and activation of a subset of T cells has/have been
CC replaced. Such a modified superantigen can optionally be used as part of
CC a conjugate with a target seeking moiety, for activating the immune
CC system to treat a mammalian disease. A pharmaceutical composition can be
CC prepared comprising a modified antibody (preferably a Fab fragment fused
CC to a peptide moiety providing activation of T cells in Vbeta specific
CC manner) in which cysteines providing for interchain cysteine linkages in
CC the native antibody have been replaced (preferably by serine residues) to
CC prohibit cysteine formation. The modified wild-type superantigen is used
CC for treating cancer, viral infections, parasitic infestations and
CC autoimmune disease. The modified wild type superantigen has a lower
CC immunogenicity and reactivity with neutralising antibodies and has fewer
CC side-effects when used as a drug, compared to wild type superantigen.

XX Sequence 233 AA;
Query Match 26.7%; Score 941; DB 18; Length 233;
Best Local Similarity 76.0%; Pred. No. 5,1e-49;
Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

QY 226 SEKSENEKDLRKSELOQTALGNLKOIYYNSKAITSSSEKSADOFNTLLPKGFTG 285
DB 1 SEKSENEKDLRKSELOQTALGNLKOIYYNEKAKTENKESHDOFLOHTILFKGFTD 60
QY 286 HPWYNDLLVLDGTAATSEYEGSSVDLYGAYGYQCAGTGNKTCMYGGVTLHDNNRLT 345
DB 61 HSWYNDLLVDFDSKDIDVKYKGVLDLYGAYGYQCAGTGNKTCMYGGVTLHDNNRLT 120
QY 346 EEKVPINLWIDGKQTTPIDKVTSSKEVTVQELDQARHYLHGKFGLYNSDSFGGKVQ 405
DB 1 SEKSENEKDLRKSELOQTALGNLKOIYYNSKAITSSSEKSADOFNTLLPKGFTG 285

DB 121 EEKVPINLWLDGKQNTVPLETVTNKNKNTVQELDQARRYLQEKYNLYNSDVFQGVQ 180
QY 406 RGLIVFHSSEGSTSYDLFDAQGOVPTDLLRIYRDNTTISSTLSISLYLYTT 458
DB 181 RGLIVFHTSTEPSVNDYLFAGQGOYSNTLLRIYRDNTINSNMHIDIYLYTS 233
RESULT 50
AAB67338
ID AAB67338 standard; peptide; 233 AA.
XX
AC AAB67338;
XX
DT 23-APR-2001 (first entry)
XX
XX Staphylococcus aureus enterotoxin A protein.
XX
KW Tumour; cancer; immune; enterotoxin.
XX
OS Staphylococcus aureus.
PN US6180097-B1.
XX
PD 30-JAN-2001.
XX
XX 30-OCT-1998; 98US-0183437.
XX
XX 31-JAN-1994; 94US-0189424.
PR
PR 19-JUN-1995; 95US-0491746.
PR
PR 03-OCT-1989; 89US-0416530.
PR
PR 17-JAN-1990; 90US-0466577.
PR
PR 17-JAN-1991; 91WO-US00342.
PR
PR 01-JUN-1992; 92US-0891718.
PR
PR 02-MAR-1993; 93US-0025144.
XX
XX (TERM/) TERMAN D S.

XX Terman DS;
XX WPI; 2001-1586657/16.
XX Tumor cell capable of stimulating antitumor immune reactivity in vitro
PT or in vivo comprises exogenous nucleic acids encoding a superantigen
PT and a costimulatory molecule -
XX
XX Disclosure; Fig 2; 16pp; English.
XX The present invention relates to a tumour cell capable of stimulating
CC antitumor immune reactivity in vitro or in vivo contains and
CC expresses an exogenous nucleic acid molecule encoding a superantigen
CC or its active fragment and an exogenous nucleic acid molecule
CC encoding a costimulatory molecule that activates T cells in
CC conjunction with an antigenic stimulus. The invention may be used
CC for cancer therapy by stimulating an anticancer immune response
CC in vivo or ex vivo.

XX Sequence 233 AA;
Query Match 26.7%; Score 941; DB 22; Length 233;
Best Local Similarity 76.0%; Pred. No. 5,1e-49;
Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

QY 226 SEKSENEKDLRKSELOQTALGNLKOIYYNSKAITSSSEKSADOFNTLLPKGFTG 285
DB 1 SEKSENEKDLRKSELOQTALGNLKOIYYNEKAKTENKESHDOFLOHTILFKGFTD 60
QY 286 HPWYNDLLVLDGTAATSEYEGSSVDLYGAYGYQCAGTGNKTCMYGGVTLHDNNRLT 345
DB 61 HSWYNDLLVDFDSKDIDVKYKGVLDLYGAYGYQCAGTGNKTCMYGGVTLHDNNRLT 120
QY 346 EEKVPINLWIDGKQTTPIDKVTSSKEVTVQELDQARHYLHGKFGLYNSDSFGGKVQ 405
DB 121 EEKVPINLWLDGKQNTVPLETVTNKNKNTVQELDQARRYLQEKYNLYNSDVFQGVQ 180

Qy 406 RGLVPHSSEGSTVSYDLEDAQQYPTLLRIYRDNTTISLSISLYLYTT 458
| | | | | : | : | | | | | : | | | | | | | | | : | : | | | :
Db 181 RGLVPHSTEPSVNYDLFGAQQYSNTLLRIYRONKTINSENMHIDIYLYTS 233

Search completed: January 8, 2003, 11:56:50
Job time : 61.5077 secs

THIS PAGE BLANK (USPTO)


```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-360-107A-122

Query Match 31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSEKXSADQFLNTLLFKGFFTG 285
Db 25 SEKSEINEKDLRKKSELORNALSRLQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
QY 286 HPWYNLLVDLGAATAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
Db 85 HPWYNLLVDLGSKDATNKYKGVLDYGYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKVPINLWIDGKQTPVDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 405
Db 145 EEKVPINLWIDGKQTPVDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 204
QY 406 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLIYRDNTTISSTSLSLYLYTT 458
Db 205 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLIYRDNKTINSENLHIDLXYLYTT 257

RESULT 4
US-08-484-223B-112
; Sequence 112, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/484,223B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-484-223B-112

Query Match 31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSEKXSADQFLNTLLFKGFFTG 285
Db 25 SEKSEINEKDLRKKSELORNALSRLQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
QY 286 HPWYNLLVDLGAATAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
Db 85 HPWYNLLVDLGSKDATNKYKGVLDYGYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKVPINLWIDGKQTPVDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 405
Db 145 EEKVPINLWIDGKQTPVDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 204
QY 406 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLIYRDNTTISSTSLSLYLYTT 458
Db 205 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLIYRDNKTINSENLHIDLXYLYTT 257

RESULT 5
US-08-919-597-112
; Sequence 112, Application US/08919597
; Patent No. 6054265
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/919,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US 08/470,896
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
```

REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-020
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-919-597-112

Query Match 31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;
QY 226 SEKSEINEKDLRKSEKLGQTALGNLKOIYYNYSKAITSSSEKSADQFLNTLLFKGFFTG 285
DB 25 SEKSEINEKDLRKSEKLGQTALGNLKOIYYNYSKAITSSSEKSADQFLNTLLFKGFFTG 84
QY 286 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 345
DB 85 HPWYNDLLVDLGSKDATNKYKGVLDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKYPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 405
DB 145 EEKYPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 406 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTSLSLSLYLYTT 458
DB 205 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTSLSLSLYLYTT 257

RESULT 6
US-08-475-668A-112
Sequence 112, Application US/08475668A
Patent No 6060065
GENERAL INFORMATION:
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 211
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,668A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-026
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 112:

SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-475-668A-112
Query Match 31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;
QY 226 SEKSEINEKDLRKSEKLGQTALGNLKOIYYNYSKAITSSSEKSADQFLNTLLFKGFFTG 285
DB 25 SEKSEINEKDLRKSEKLGQTALGNLKOIYYNYSKAITSSSEKSADQFLNTLLFKGFFTG 84
QY 286 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 345
DB 85 HPWYNDLLVDLGSKDATNKYKGVLDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKYPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 405
DB 145 EEKYPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 406 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTSLSLSLYLYTT 458
DB 205 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTSLSLSLYLYTT 257

RESULT 7
US-08-485-551A-112
Sequence 112, Application US/08485551A
Patent No 6068973
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 211
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,551A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-023
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid

```

; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-551A-112

Query Match          31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKKSELOQTALGNLKQIYYNSKAITSEKSDOFLNTWLLFKGFFTG 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 SEKSEINEKDLRKKSELQNALSNLRQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 286 HPWYNLLVDLGSTATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 HPWYNLLVDLGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 346 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 405
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 145 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 406 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLLRIYRDNTTISSTLSLSLYLYTT 458
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLLRIYRDNKTINSENLHIDLILYLYTT 257
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-08-471-913A-112
; Sequence 112, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/471,913A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
```

```

US-08-471-913A-112

Query Match          31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKKSELOQTALGNLKQIYYNSKAITSEKSDOFLNTWLLFKGFFTG 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 SEKSEINEKDLRKKSELQNALSNLRQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 286 HPWYNLLVDLGSTATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 HPWYNLLVDLGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 346 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 405
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 145 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 406 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLLRIYRDNTTISSTLSLSLYLYTT 458
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLLRIYRDNKTINSENLHIDLILYLYTT 257
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-08-485-264A-112
; Sequence 112, Application US/08485264A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/485,264A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-264A-112

Query Match          31.4%; Score 1107; DB 4; Length 257;
```

		Qy	226	SEKSEEINEKDLRKKSELQGTAIGNLKOIYYNYSKAITSSEKSADOPLTNTLLFKGFTTG	2895
		Dd	25	SEKSEEINEKDLRKKSELQRNALSNLROIYYNEKAITENKESDDOPLENTLLFKGFTTG	84
		Qy	286	HPWYNDLVLDLGSTAAATSEVEGGSSVDLYGAYGYQCAGGTTPNKTCACMGVTLHDNNRLT	3451
		Dd	85	HPWYNDLVLDLGSDATANKYKGKRVDLYGAYGYQCAGGTTPNKTCACMGVTLHDNNRLT	1441
		Qy	346	EKKVPINLWDGKTTPIDVKYTSKEVTVOELDQAARHYLHGKFGLNSDSFGGKVQ	405
		Dd	145	EKKVPINLWDGKTTPIDVKYTSKEVTVOELDQAARHYLHGKFGLNSDSFGGKVQ	204
		Qy	406	RGLVFHSSEGSVSYDLFDAQGYPTDLLRIYRDNNTTISTSLISLYLYTT	458
		Dd	205	RGLVFHSSGESVSVDLPDAOGYPDTLLRIYRDKNKITENSENHIDLILYTT	257

```

RESULT 11
US-08-353-400-33
; Sequence 33, Application US/08353400
; Patent No. 5665357
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 37
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,400

```

```

?      FILING DATE: 03-DEC-1993
?
?      PRIOR APPLICATION DATA:
?      APPLICATION NUMBER: GB 9411089.7
?      FILING DATE: 03-JUN-1994
?      INFORMATION FOR SEQ ID NO: 33:
?      SEQUENCE CHARACTERISTICS:
?          LENGTH: 445 amino acids
?          TYPE: amino acid
?          STRANDEDNESS: single
?          TOPOLOGY: linear
?          MOLECULE TYPE: protein
?      US-08-353-400-33

```

Qy	1	EVOLQSQSPDLVAPGASVKISCRASGYSFYGYGMHVKQSPGKGLWIGRINPNNGVTLY	60
Db	1	QVQLQQQGAELVPGASVQUSCRASGYFTFGYIHHWKQRPQGGLEWIGEVNPGSTGRSDY	60
Qy	61	NOQFKDKATLTVDKSSITAYMELRSLTSDSVAVYCARSTMI-TNYMNDYWGOGTSVTVS	119
Db	61	NEKFKKATLTVDKSSITAYMQLSSLTSDSVAVYCARERAYCYDAMDYWGOGTSVTVS	120
Qy	120	SAKTTTPSVVPLAPGSAAQNSMWTGLCLVKGYFPEPVTVTWNSGSLSSGVHFTFPAVLQS	179
Db	121	SAKTTTPSVVPLAPGSAAQNSMWTGLCLVKGYFPEPVTVTWNSGSLSSGVHFTFPAVLQS	180
Qy	180	DLYTLSSSVVTPSSTWPSSEVTCNVAHPASSTKVDKKIIVPRDSGGPSEKSEENEXDLRK	239
Db	181	DLYTLSSSVVTPSSTWPSSEVTCNVAHPASSTKVDKKIIVPRDCG-----	224
Qy	240	KSEIQGTALGNLKOIYYNYSKATISSEKSAQDFLTNTLFPKGFPTGHPWYNDLVLDLGST	299
Db	225	-CKPCICITVPSVSSVFTFPKK-----PKDVLITPL-----	253

```
Qy 300 AATSEYEGSSVDLYGAYYGOCAGGTPNKTACMYGGVTLHDNNRLTEBKKVPINLWIDGK 359
Db 254 -----TP-KVTCVVVDIS-----KDDPEVQFSNFDV-- 278
Qy 360 QTTVPIDKVKTSKEV-----TVQELDLQARHYLHCK-FGL-YNSDSFGGKVQORGLIV 410
Db 279 --DVEHTAQTQPREEQNFSTRSVSELPIMHQDMLNGKEFKRCRVNSAAFPAPIEK---- 332
Qy 411 FHSSEGSTSVSYDLFDAQGYPDTLRLIRYDNTTISSTLSISLYLYTTSIVMTQPTSL 470
Db 333 -----TISK----- 337
Qy 471 VSAGDRVITCKASOSVNDVANYQQKPGQSPKLLISYSSRYAGVDPDRFSGSGYGTDF 530
Db 338 ----- 337
Qy 531 LTISVQAEDAAYVFCQDYNPPFTGGGTLKLEIKRADAAPTSTVFPPSSBOLTSGGASV 590
Db 338 -----KGRPKAPOVYTIPPKEQMAKDKVSL 363
Qy 591 VCFNLNFKPDINVKWKIDGSEQRQNGVLNSWTDQDSKDSYMSSTLTLTDEYERHNSY 650
Db 364 TCMITDFFPEDITVEMQWNGQPAEN-YKNTQPIMDT-DGSYFVYSKLVQKSNWAGNTF 421
Qy 651 TCEATHK 657
Db 422 TCSVLHE 428

RESULT 12
US-08-353-400-36
; Sequence 36, Application US/08353400
; Patent No. 5665357
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 37
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353.400
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9324819.3
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9411089.7
; FILING DATE: 03-JUN-1994
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 464 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-353-400-36
```

```
Query Match 30.1%; Score 1059.5; DB 1; Length 464;
Best Local Similarity 37.5%; Pred. No. 2.2e-63;
Matches 250; Conservative 61; Mismatches 107; Indels 249; Gaps 14;

Qy 1 EVOLQSQPDLYKPCASVKISKASGYFTGYHHWVKQSGKGLWIGRINPNNGVTLY 60
Db 20 QVQLQQPAELVPGASVOLSKASGYFTGYIHWKQRCQGLWIGVNPSTGRSDY 79
Qy 61 NOKFKDKATLVDSKSTTAYMELRSLTSEDSAVYYCARSTMI-TNYVMDYQGQTSVTS 119
Db 80 NEKFNKATLVDSKSTTAYMQLSLTSEDSAVYYCARERAYGDAMDYQGQTSVTS 139
```

```
Qy 120 SAKTTPPSVYPLAPGSAATNSMVTGLCLVKGYFPEPVVTVTWSGSLSSGVTHTFPAVLQS 179
Db 140 SAKTTPPSVYPLAPGSAATNSMVTGLCLVKGYFPEPVVTVTWSGSLSSGVTHTFPAVLQS 199
Qy 180 DLYTLSSSVTPSSWTPSETVTCNVAHPASSTKVDDKLVPRDSGPGSPSEKSEINEKDLRK 239
Db 200 DLYTLSSSVTPSSWTPSETVTCNVAHPASSTKVDDKLVPRDCG----- 243
Qy 240 KSELQGTALGNLKOIYYYNKAITSSSEKSAQOFLTNTLLFKGFTGHPWYNDLLVDLGS 299
Db 244 -CKPCICIVPESSVFIPPK-----PKDVLITIL----- 272
Qy 300 AATSEYEGSSVDLYGAYYGOCAGGTPNKTACMYGGVTLHDNNRLTEBKKVPINLWIDGK 359
Db 273 -----TP-KVTCVVVDIS-----KDDPEVQFSNFDV-- 297
Qy 360 QTTVPIDKVKTSKEV-----TVQELDLQARHYLHCK-FGL-YNSDSFGGKVQORGLIV 410
Db 298 --DVEHTAQTQPREEQNFSTRSVSELPIMHQDMLNGKEFKRCRVNSAAFPAPIEK---- 351
Qy 411 FHSSEGSTSVSYDLFDAQGYPDTLRLIRYDNTTISSTLSISLYLYTTSIVMTQPTSL 470
Db 352 -----TISK----- 356
Qy 471 VSAGDRVITCKASOSVNDVANYQQKPGQSPKLLISYSSRYAGVDPDRFSGSGYGTDF 530
Db 357 ----- 356
Qy 531 LTISVQAEDAAYVFCQDYNPPFTGGGTLKLEIKRADAAPTSTVFPPSSBOLTSGGASV 590
Db 357 -----KGRPKAPOVYTIPPKEQMAKDKVSL 382
Qy 591 VCFNLNFKPDINVKWKIDGSEQRQNGVLNSWTDQDSKDSYMSSTLTLTDEYERHNSY 650
Db 383 TCMITDFFPEDITVEMQWNGQPAEN-YKNTQPIMDT-DGSYFVYSKLVQKSNWAGNTF 440
Qy 651 TCEATHK 657
Db 441 TCSVLHE 447

RESULT 13
US-08-896-933-24
; Sequence 24, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896.933
; CURRENT FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-24

Query Match 29.5%; Score 1040; DB 4; Length 226;
Best Local Similarity 85.2%; Pred. No. 1.8e-62;
Matches 196; Conservative 13; Mismatches 17; Indels 4; Gaps 1;

Qy 229 SEETNEKDLRKSELOQTALGNLKOIYYYNKAITSSSEKSAQOFLTNTLLFKGFTGHPW 288
Db 1 SEETNEKDLRKSELOQTALGNLKOIYYYNKAITSSSEKSAQOFLTNTLLFKGFTGHPW 60
Qy 289 YNDLLVDLGSATAATSEYEGSSVDLYGAYYGOCAGGTPNKTACMYGGVTLHDNNRLTEK 348
Db 61 YNDLLVDLGSKDATNKYKGVVDLYGAYYGOCAGGTPNKTACMYGGVTLHDNNRLTEE- 119
```

Qy 349 KVPINLWIDGKQTTVPIDKVTSSKEVTVDLQARHLYHGKFGLYNSDSFGGKVQRGL 408
Db 120 ---VBKWIDGKQTTVPIDKVTSSKEVTVDLQARHLYHGKFGLYNSDSFGGKVQRGL 176
Qy 409 IVFHSSEGSTVSYDLFDAGQGYPTDTPLLRIYRDNNTTISSTLSISLYLTT 458
Db 177 IVFHSSEGSTVSYDLFDAGQGYPTDTPLLRIYRDNNTTISSTLSISLYLTT 226

RESULT 14

US-09-314-235-24
; Sequence 24, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; CURRENT FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416,530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-24

Query Match 29.5%; Score 1040; DB 4; Length 226;
Best Local Similarity 85.2%; Pred. No. 1.8e-62;
Matches 196; Conservative 13; Mismatches 17; Indels 4; Gaps 1;

Qy 229 SEINEKDLRKSELOGTALGNLKOIYYNSKAITSSSEKSADQELTNTLLFKGFTGHPW 288
Db 1 SEINEKDLRKSELOGTALGNLKOIYYNSKAITENKESDDQELNTLLFKGFTGHPW 60
Qy 289 YNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHDNNRLTTEK 348
Db 61 YNDLLVDKGSADATNKYKGVVDLYGAYGYQCAGGTPNKTCMYGGVTLHDNNRLTEE- 119
Qy 349 KVPINLWIDGKQTTVPIDKVTSSKEVTVDLQARHLYHGKFGLYNSDSFGGKVQRGL 408
Db 120 ---VBKWIDGKQTTVPIDKVTSSKEVTVDLQARHLYHGKFGLYNSDSFGGKVQRGL 176
Qy 409 IVFHSSEGSTVSYDLFDAGQGYPTDTPLLRIYRDNNTTISSTLSISLYLTT 458
Db 177 IVFHSSEGSTVSYDLFDAGQGYPTDTPLLRIYRDNNTTISSTLSISLYLTT 226

RESULT 15

US-09-485-737B-90
; Sequence 90, Application US/09485737B
; Patent No. 6350860
; GENERAL INFORMATION:
; APPLICANT: Buysse, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC SHOCK,
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015
; CURRENT APPLICATION NUMBER: US/09/485,737B

; CURRENT FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
US-09-485-737B-90

Query Match 27.8%; Score 980.5; DB 4; Length 711;
Best Local Similarity 35.3%; Pred. No. 7.6e-58;
Matches 254; Conservative 103; Mismatches 180; Indels 183; Gaps 23;

Qy 1 EVOIQSGPDLVKFGASVKISKASGYSTGYMHWKQSPGKLEWIGRINPNNGVTLY 60
Db 21 QVQLVQSGSELKPCASVKISKASGYSTGYMHWKQAPCQGLKMWGINTYTGSTY 80
Qy 61 NQKFKDAFLTVDKSSTPAYMELRSLTSEDSAVYVCARSTMITNVMYDYGOGTSVTYSS 120
Db 81 VDDFKGRVFSLDTSVSAAYLQISLKAEDATYFCARGF---YAMDYWGOGTIVTVSS 137
Qy 121 AKTTPPSVYPLAGSAAQTNSMVTIGCLVKGYFPEPVTWNSGSLSSGVHTFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAAALGCLVKLVKDYFPEPVTWNSGALTSGVHTFPAVLQSS 197
Qy 181 -LYTLSSSVTPSPSTWSETVTCNVAPASSTKVDKKIVPRDS-----GG 224
Db 198 GLYSLSSVTVPSLSLGTQTYICNVNHPKPSNTKVDKRVKPKSCDKTHTCPCPAPELGG 257
Qy 225 PS-----EKSEI-----NEKDLRKSELOGTALGNLK---QIYYN 258
Db 258 PSVFLFPKPKDMLMISRTPEVTCVVDVSHEDPEVKFNWYDGVVEVHNAKTPREEQYN 317
Qy 259 S-----KAITSS-EK-----SADQ 271
Db 318 STYRVSVSLTVLHQDWLNGKEYCKKVSNNKALPASTIEKTSKAKGQPREPQVYVTLPPSREE 377
Qy 272 FLTN-----TLFKGFTGHPWYNDLLVDLGSTA-ATSEYEGSS--VDLYGAYGYQCAGG 324
Db 378 MTKNQVSLTCLVKGFYF-----SDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSKL-- 430
Qy 325 TPNNKTACHYGGV---TLHD--NNRLTTEK-----KVPINLWIDGKQTTVPIDKV 368
Db 431 TVDKSRWQOGNVFSCSVMHREALHNYTOKLSLSPGKLGSGSQVLQVSGSELKPKGASV 490
Qy 369 KTSKKEVTVDLQARHLYHGKFGLYNSDSFGGKVQRGLIVFHSSEGSTVSYDLFDAGQ 428
Db 491 KISCK-----ASCYTFDYGMNWKVQAQGLKMWGINTYTGSTYVD--DFKG 538
Qy 429 QYPTDTPLLRIYRDNNTTISSTLSIS-----LYLYT----- 457
Db 539 RF-----VFLSDTSVSAAYLQISLKAEDATYFCARRGFYAMDYWGOGTIVTVSSGGG 592
Qy 458 -----TSIVMTQTPPTSLVSGADRVTTITCASQSVSNDVAVYQKPGSPKLLI 506
Db 593 GSGGGGGGGSDIVLTQSPATMSASGGERVTLTCSASSISY-MFWTHQRPGSGPRLLI 651
Qy 507 SYTSSRYAGVPDRFSGGCGYGTFTLTSSVQAEDAAYFCQQDYNSPPTFGGKTKLEIR 566
Db 652 YDTSNLAGVPAFSGGSGGTYSLTISRMEPEDFATYFCHQSSSYPTFGGKTKLEIR 711

RESULT 16
us-08-116-247-7
; Sequence 7, Application US/08116247

Query Match	27.4%	Score	966;	DB 2;	Length	468;			
Best Local Similarity	34.9%	Pred.	No. 4.3e-57;						
Matches	241;	Conservative	56;	Mismatches	126;	Indels	268;	Gaps	14;
Qy	1	EVOLQQSGPDLVKPGASVKISCKASGYSFTGYMHWVKQSPGKLEWIGTRINPNNGVLY	60						
Db	20	QVQLQQSGAEELARPGASVKKSCASGYTFRTYTHHWVKQRPGQGLEWIGLYINPSRGYTN	79						
Qy	61	NQKPKKATILTVKSSSTTAYMELRSLTSEDSAVYYCARSTMTINVMYDYGQGTSTVTVSS	120						
Db	80	NQKFKDKATILTTDKSSSTAYMQLSSLTSEDSAVYYCAR--YYDDHYCYLDYWGQGTTLTVSS	138						
Qy	121	AKTTPPSVYPLAPGSAQAOTNSMTLGLCLVKGYFPEPVTLTVNWSGSLSGSVTFPAVLQSD	180						
Db	139	AKTTAPSVYPLAPVCGGDTTGGSVTLGGLCLVKGYFPEPVTLTVNWSGSLSGSVTFPAVLQSD	198						
Qy	181	LYTLSSSVTPSPSTWSPSETVCNVAHPASSTKVDKIVPRDSGGSPSEKSEINEKDLRKK	240						
Db	199	LYTLSSSVTPSTWSPSQSITCNVAHPASSTKVDKKIEPR--GPTIKPCP-----	246						
Qy	241	SELQGTALGNLKQIYYYNKRAITSSEKSAOFLTNTLLFKGFFTHGHPWYNOLLVDLGSTA	300						
Db	247	-----	246						
Qy	301	ATSEYEGSSVDLYCAYYGYOCAGGTPNKTAACMGCVTLHDNNRLTEKKVPINLWIDGKQ	360						
Db	247	-----PCKCPAPN---LLGGPSVF-----	262						
Qy	361	TTVPIDKVKTSKKEVTVQVELDQARHYLHGKFGLYNSDSFGGKVGORGLIVFHSSEGSTVS	420						
Db	263	-----IFPPKTKDVLMI---SLSPVLT	281						

Db 2 ESGTELAKPGASVKMSCKASGYTFTSWIHWKORPGOGLEWIGYINPSTDYTYEYIQKFK 61
Qy 66 DKATLTVDKSTTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGOGTSVTVSSAKTTP 125
Db 62 DKATLTADKSSSTAYMQLSLTSDSAVYYCV-----MKDYWGOGTSVTVSSAKTTP 113
Qy 126 PSVYPLAPGSAOATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSDLYTLS 185
Db 114 PSVYPLAPGSAOATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSDLYTLS 173
Qy 186 SSVTPSPSTPSETVTCNVAHPASSTKVKKIVPRD 221
Db 174 SSVTPSPSTPSETVTCNVAHPASSTKVKKIVPRD 209

RESULT 18
US-08-303-569B-7
: Sequence 7, Application US/08303569B
: Patent No. 5859205
: GENERAL INFORMATION:
: APPLICANT: Adair, John R.
: APPLICANT: Athwal, Diljeet S.
: APPLICANT: Emtege, John S.
: TITLE OF INVENTION: Humanised Antibodies
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5859205ris
: STREET: One Liberty Place - 46th Floor
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/303.569B
: FILING DATE: 07-SEP-1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Trujillo, Doreen Yanko
: REGISTRATION NUMBER: 35,719
: REFERENCE/DOCKET NUMBER: CARP-0032
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 568-3100
: TELEFAX: (215) 568-3439
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 468 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-303-569B-7

Query Match 27.2%; Score 957; DB 2; Length 468;
Best Local Similarity 34.7%; Pred. No. 1,6e-56;
Matches 240; Conservative 56; Mismatches 127; Indels 268; Gaps 14;

Qy 1 EVOLQSQGDLVKPGASVKISCKASGYFTGYMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db 20 QVOLQSQGAELARPGASVKMSCKASGYFTTRYTMHWKQSPCKGLEWIGYINPSRGTYNY 79
Qy 61 NQKFKDKATLTVDKSSSTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGOGTSVTVSS 120
Db 80 NQKFKDKATLTVDKSSSTAYMQLSLTSDSAVYYCAR--YYDDHYCLDYWGOGTTLTVSS 138
Qy 121 AKTTPSPVYPLAPGSAOATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSD 180
Db 139 AKTTPSPVYPLAPGCDTGTSSVTLGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSD 198
Qy 181 LYTSSSVTPSPSTPSETVTCNVAHPASSTKVKKIVPRDSDSGPSEKSEINEKDLRKK 240

Db 199 LYTSSSVTPSPSTPSETVTCNVAHPASSTKVKKIEPR---GPTIKPCP----- 246
Qy 241 SELQGTALGNLQIYYNKAITSSEKSAQDLFTNTLLFGFTGHPWYNDLLVDLAGSTA 300
Db 247 ----- 246
Qy 301 ATSEYEGSSVDLYGAYYGQCAGGTGPKNTACMGVTLHDNNRLTEKKKVPINLWIDGKQ 360
Db 247 -----PCKCPAPN-----LLGGPSVF----- 262
Qy 361 TTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDFGKVGQRLIVPHSSEGSTVS 420
Db 263 -----IFPKIKDVLMI---SLSPIVT 281
Qy 421 YDLFDAQGYQDPTLLRIYRDNNTTISSTLSLSLYTTSIVMTQTPTSLTLLVSAGDRVTIT 480
Db 282 CVVDVDSDDPD-----VOISWFEVNNVEVHTAQQT----- 312
Qy 481 CKASQSVSNDAVYQOQPGQSPKLLISYTSRYAGVDPDRFGSGGYGDTFTLTISVQAE 540
Db 313 -----HREDYNSTLRV----- 324
Qy 541 AAYVFCQDYNPPTFGGCTKLEIKRAD-----AAPTYSIFPPSSEOLT 584
Db 325 SALPIQHDWMSGKEF-----KCKVNNKDLPAPIERTISKPGSVRAPQVYVLPPEEEMT 380
Qy 585 SGGASVVCFLANFYPKDINVKWKIDGSRQN-----GYLNSWTDQDSKDSYSSMSLTLL 639
Db 381 KKQVTLTCMVTDFMPEDIYVETNNGKTELAKYKNTPEVLD-----DGSYFMYSKLRV 433
Qy 640 TKDEYERHNSYTCEATHK-TSTSPIVKSFN 669
Db 434 EKKNNVRSYSCSVVHEGLHNNHHTKSF 464

RESULT 19
PCT-US94-14106-51
: Sequence 51, Application PC/TUS9414106
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Process for Generating Specific Antibodies
: NUMBER OF SEQUENCES: 61
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII (text)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/14106
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 51:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 223 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
PCT-US94-14106-51

Query Match 27.1%; Score 953.5; DB 5; Length 223;
Best Local Similarity 83.0%; Pred. No. 1.1e-56;
Matches 185; Conservative 12; Mismatches 23; Indels 3; Gaps 2;

Qy 1 EVOLQSQGDLVKPGASVKISCKASGYFTGYMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db 1 EVOLQSQGPELMMPGASVKISCKATGYTLSSYWLKQSPCKGLEWIGELIFLGGSAHY 60
Qy 61 NQKFKDKATLTVDKSSSTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGOGTSVTV 118
Db 61 NEKFKGKATFTVDTSSNTAYMQLSLTSDSAVYYCARGD-YGNYGDYDYWGOGTTLTV 119
Qy 119 SSAKTTTPSPVYPLAPGSAOATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQ 178

Db 120 SSAKTPPSVYPLAPGSAQTNSMYTLCCLVKGYPPEVTVTNWSGLSSGSHVTFPAVLQ 179
Qy 179 SLYTLSSSVTPSPSEVTVCNVAHPASSTKVKDKKIVPRD 221
Db 180 SLYTLSSSVTPSPSEVTVCNVAHPASSTKVKDKKIVPRD 222

RESULT 20
US-08-486-099-113
: Sequence 113, Application US/08486099
: Patent No. 6013263
: GENERAL INFORMATION:
: APPLICANT: Bolognesi, Dani P.
: APPLICANT: Matthews, Thomas J.
: APPLICANT: Wild, Carl T.
: APPLICANT: Barney, Shawn O.
: APPLICANT: Lambert, Dennis M.
: APPLICANT: Petteway, Stephen R.
: APPLICANT: Langlois, Alphonse J.
: TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
: TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
: TITLE OF INVENTION: B VIRUS TRANSMISSION
: NUMBER OF SEQUENCES: 209
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/486,099
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7872-031
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 113:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 257 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-486-099-113

Query Match 26.9%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 226 SEKSEINEKDLRKSELQGTALGNLKOIYYNYSKAITSSSEKSADQFLTNLLFKGFFTG 285
Db 25 SEKSEINEKDLRKSELQGTALGNLKOIYYNYSKAITSSSEKSADQFLTNLLFKGFFTG 84

Qy 286 HPWNLDLLVGLGTAATSEYEGSSVDLYGAYGYQCAGTGNKTCACMYGGVTLHDNNRLT 345
Db 85 HSWNDLLVDFDSDIVDKYKGGKVDLYGAYGYQCAGTGNKTCACMYGGVTLHDNNRLT 144

Qy 346 EEKVPINLWIDGKQTTVPIDKVKTSKEVTQVELDQARHYLHCKFGLYNSDSFGGKVQ 405
Db 145 EEKVPINLWIDGKQNTVPLETVTKNKNVTQVELDQARRYLOEKYNLYNSDVFQGVQ 204

Qy 406 RGLIVFHSSEGSTVSYDLFDAQGYQYPTDRLRIYRDNFTTISSTLSLSISLYTT 458
Db 205 RGLIVFHTSTPSVNYDLFGAQGYSNFTLLRIYRDNKTINSENHIDIYLYTS 257

RESULT 21
US-08-360-107A-123
: Sequence 123, Application US/08360107A
: Patent No. 6017536
: GENERAL INFORMATION:
: APPLICANT: Bolognesi, Dani P.
: APPLICANT: Matthews, Thomas J.
: APPLICANT: Wild, Carl T.
: APPLICANT: Barney, Shawn O.
: APPLICANT: Lambert, Dennis M.
: APPLICANT: Petteway, Stephen R.
: APPLICANT: Langlois, Alphonse J.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
: TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
: TITLE OF INVENTION: TRANSMISSION
: NUMBER OF SEQUENCES: 149
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/360,107A
: FILING DATE: 20-DEC-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7872-013
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 123:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 257 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-360-107A-123

Query Match 26.9%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 226 SEKSEINEKDLRKSELQGTALGNLKOIYYNYSKAITSSSEKSADQFLTNLLFKGFFTG 285
Db 25 SEKSEINEKDLRKSELQGTALGNLKOIYYNYSKAITSSSEKSADQFLTNLLFKGFFTG 84

Qy 286 HPWNLDLLVGLGTAATSEYEGSSVDLYGAYGYQCAGTGNKTCACMYGGVTLHDNNRLT 345
Db 85 HSWNDLLVDFDSDIVDKYKGGKVDLYGAYGYQCAGTGNKTCACMYGGVTLHDNNRLT 144

Qy 346 EEKVPINLWIDGKQTTVPIDKVKTSKEVTQVELDQARHYLHCKFGLYNSDSFGGKVQ 405
Db 145 EEKVPINLWIDGKQNTVPLETVTKNKNVTQVELDQARRYLOEKYNLYNSDVFQGVQ 204

Qy 406 RGLIVFHSSEGSTVSYDLFDAQGYQYPTDRLRIYRDNFTTISSTLSLSISLYTT 458
Db 205 RGLIVFHTSTPSVNYDLFGAQGYSNFTLLRIYRDNKTINSENHIDIYLYTS 257

RESULT 22

US-08-484-223B-113
; Sequence 113, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-223B-113

Query Match 26.9%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy	226	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSSEKSADOFLTNTLLFKGFFTG	285
Db	25	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSSEKSADOFLTNTLLFKGFFTD	84
Qy	286	HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYYGQACGTPNKTACMYGGVTLHDNNRLT	345
Db	85	HSWYNDLLVDFDSKDIVDKYKGVLDLYGAYYGQACGTPNKTACMYGGVTLHDNNRLT	144
Qy	346	EEKKVPINLWIDGKQTTPIDKVKTSKKEVTVQVELDQARHYLHGKFGLYNSDFGKQV	405
Db	145	EEKKVPINLWIDGKQNTVPLETVTKNKNVTVQVELDQARRYLQEKYNLYNSDFDGKQV	204
Qy	406	RGLIVFHSSEGSTSYDLFDAQGYPDFTLLRIYRDNNTTISSTSLISLYLYTT	458
Db	205	RGLIVFHTSTEPSVNDYDFGAGQYSNTLLRIYRDNNTTINSENHIDIYLYTS	257

RESULT 23

US-08-919-597-113
; Sequence 113, Application US/08919597
; Patent No. 6054265
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/470,896
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-919-597-113

Query Match 26.9%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy	226	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSSEKSADOFLTNTLLFKGFFTG	285
Db	25	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSSEKSADOFLTNTLLFKGFFTD	84
Qy	286	HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYYGQACGTPNKTACMYGGVTLHDNNRLT	345
Db	85	HSWYNDLLVDFDSKDIVDKYKGVLDLYGAYYGQACGTPNKTACMYGGVTLHDNNRLT	144
Qy	346	EEKKVPINLWIDGKQTTPIDKVKTSKKEVTVQVELDQARHYLHGKFGLYNSDFGKQV	405
Db	145	EEKKVPINLWIDGKQNTVPLETVTKNKNVTVQVELDQARRYLQEKYNLYNSDFDGKQV	204
Qy	406	RGLIVFHSSEGSTSYDLFDAQGYPDFTLLRIYRDNNTTISSTSLISLYLYTT	458
Db	205	RGLIVFHTSTEPSVNDYDFGAGQYSNTLLRIYRDNNTTINSENHIDIYLYTS	257

RESULT 24

ADDRESSEE: Pennle & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/474.349A
APPLICATION NUMBER: US/08/474.349A
FILING DATE: 07-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-024
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-474-349A-113

Query Match 26.9%; Score 948; DB 4; Length 257;
Best Local Similarity 76.4%; Pred. No. 3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSDAQFILTNTLLFGKFFTG 285
DB 25 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSDAQFILTNTLLFGKFFTG 84
QY 286 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGTGNKTACMYGGVTLHDNNRLT 345
DB 85 HSWYNDLLVDFDSKDIDVYKGGKVDLYGAYGYOCAGTGNKTACMYGGVTLHDNNRLT 144
QY 346 EEKVPINLWDGKQTVPIDKVKTSKREVTQVELDLQARHYLHGKFGLYNSDSFGKVKQ 405
DB 145 EEKVPINLWDGKQTVPLETVTKNKNVTVQVELDLQARHYLQEKYLYNSDVFQKVKQ 204
QY 406 RGLIVFHSSEGSTVSYDLFDAGQGVDPDRLRIYRDNNTTSSLSLSLYLYTT 458
DB 205 RGLIVFHTSTEPSVNYDLFGAQGGYSNTLLRIYRDNKTNSNMHIDIYLYTS 257

RESULT 29
US-08-353-400-37
Sequence 37, Application US/08353400
Patent No. 5665357
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 37
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353.400
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9324819.3
FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9411089.7
FILING DATE: 03-JUN-1994
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-400-37

Query Match 26.8%; Score 944.5; DB 1; Length 239;
Best Local Similarity 78.1%; Pred. No. 4.8e-56;
Matches 185; Conservative 22; Mismatches 19; Indels 11; Gaps 3;

QY 445 SSTSLSISLYLYTT----SIVMTQTPTSLLYSAGDRVTITCKASOSVSD-----VAVY 494
DB 3 SOAQVLILLLLWVSGTGDIVMSQSPSSSLAVSAGEKVTMSCKSSQSLNSRTRKNYLAWY 62
QY 495 OOKPGOSPCKLLISYTSRYAGVPDRFSGSGYGTDFTLTIISSVQAEADAAYVFCQOQDYNSSPP 554
DB 63 QORPGOSPCKLLIYWASTRTSGVPDRFTGSGGTDFTLTIISSVQAEADLAIYYCKQSY-TLR 121
QY 555 TFGGOTKLEIKRADAAPTVSIFPPSSEQLTSGGASVWCFLNNFYPKDINVKWKIDGSRQ 614
DB 122 TFGGOTKLEIKRADAAPTVSIFPPSSEQLTSGGASVWCFLNNFYPKDINVKWKIDGSRQ 181
QY 615 NGVLNSWTDQDSKSTYSMSSTLTLTDKDEYERHNSYTCETHKTSTSPIVKSFNRNE 671
DB 182 NGVLNSWTDQDSKSTYSMSSTLTLTDKDEYERHNSYTCETHKTSTSPIVKSFNRNE 238

RESULT 30
US-08-446-918A-4
Sequence 4, Application US/08446918A
Patent No. 5705151
GENERAL INFORMATION:
APPLICANT: Dow, Steve W.
APPLICANT: Elmslie, Robyn E.
TITLE OF INVENTION: GENE THERAPY FOR T CELL REGULATION
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446.918A
FILING DATE: 18-MAY-1995
CLASSIFICATION: 552
ATTORNEY/AGENT INFORMATION:
NAME: Kovarik, Joseph E.
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2879-29
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-918A-4

```

Query Match          26.8%; Score 944; DB 1; Length 233;
Best Local Similarity 76.3%; Pred. No. 5e-56;
Matches 177; Conservative 21; Mismatches 34; Indels 0; Gaps

QY 227 EKSEINEKDLRKSELOQTALGNLKOIYYVNSKAITSSSEKSADQFLTNTLLFKGFFTGH 286
Db 2 EKSEINEKDLRKSELOQTALGNLKOIYYVNSKAITSSSEKSADQFLTNTLLFKGFFTGH 61
QY 287 PWTNDDLLVDLGSAATSEYEGSSVDLYGAYGYQCAAGTGNKACMYGGVTLHDNNRLTE 346
Db 62 SWYNDLLVDFDSKDIDVKYGGKKVDLYGAYGYQCAAGTGNKACMYGGVTLHDNNRLTE 121
QY 347 EKKVPINLWJDGQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKGLYNSDSFGKQVQR 406
Db 122 EKKVPINLWJDGQNTVPLETVKTNKKNVTVOELDLQARHYLQEKYLNYSDFDGVKQVR 181
QY 407 GLIVFHSSEGSTVSYDLFDAQGOYPTLLRIYRDNVTTISSTLSLSLYLYTT 458
Db 182 GLIVFHTSTEPSYNYDLFGAAGOYSNTLLRIYRDNKNTINSENMHIDILYTS 233

RESULT 31
US-08-580-806-4
; Sequence 4, Application US/08580806
; Patent No. 5935568
; GENERAL INFORMATION:
; APPLICANT: Dow, Steve W.
; APPLICANT: Elmslie, Robyn E.
; APPLICANT: Potter, Terence A.
; TITLE OF INVENTION: GENE THERAPY FOR EFFECTOR CELL REGULATION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580,806
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2879-29-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PS-08-580-806-4

```

```

Qy 347 EKKVPINLWDGKQTTVPIDKVKTSKKEVTVQVELDLQARHYLHGKFGLYNSDSFGGKVQR 406
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 122 EKKVPINLWDGKQNTVPLETKVTKNKKNTVQVELDLQARRYLQEKYLNLYNSDVPDGKVQR 181
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Qy 407 GLIVFHSEGSSTSYDLDFAOGXYPDITLLRIYRDNNTTISSTLSLSLYLYTT 458
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 182 GLIVFHTSTEPSVNDYDFGAOGYSNITLLRIYRDNKTINSENMHIDIYLYTS 233
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 32
US-08-353-400-34
; Sequence 34, Application US/08353400
; Patent No. 5663357
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 37
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,400
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9324819.3
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9411089.7
; FILING DATE: 03-JUN-1994
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 219 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-353-400-34

Query Match 26.7%; Score 941.5; DB 1; Length 219;
Best Local Similarity 83.0%; Pred. No. 6.8e-56;
Matches 181; Conservative 19; Mismatches 11; Indels 7; Gaps 2;

Qy 460 IVMTQTPTSLVLSAGDRVITITCKASQSVSND-----VAVYQQRPGQSPKLLISYTSRY 513
      ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2 IVMSQSPSSLAVSAGEKVTMSCKSSQSLNSTRKNLYLAWYQQRPGQSPKLLIYWASTRT 61

Qy 514 AGVPRDRSGSGYGTDFITLTSSVQAEDAAYVFCQODYNSPPTFGGGTKLETIKRADAAPT 573
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 62 SGVPDRFTGSGSGDFTLTISVQAEDLAIYCKQSY-TLRTFGGGTKLETIKRADAAPT 120

Qy 574 SIFFPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSEKONGVLSNWTDDSKDSTYSM 633
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 SIFFPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSEKONGVLSNWTDDSKDSTYSM 180

Qy 634 SSTLTGKDEYRHNSYTCEATHKTTSTSPIVKSFNRRN 671
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 SSTLTGKDEYRHNSYTCEATHKTTSTSPIVKSFNRRN 218

RESULT 33
US-08-737-129A-6
; Sequence 6, Application US/08737129A
; Patent No. 5885816
; GENERAL INFORMATION:
; APPLICANT: Ikuo FUJII et al.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES ENANTIOSELECTIVELY
; HYDROLYSING AMINO ACID ESTER DERIVATIVES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737.129A
FILING DATE: No. 5885816ember 15, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-737-129A-6

Query Match 26.5%; Score 933; DB 2; Length 222;
Best Local Similarity 82.6%; Pred. No. 2.5e-55;
Matches 180; Conservative 12; Mismatches 24; Indels 2; Gaps 1;
QY 6 QSGPDLVKGASVKISCKASGYFTGYVMHWKQSPGKLEWIGRINPNNGVTLYNOKFK 65
DB 2 ESGPELVKGGVSTISCKASGFTTSMNVRQRPQGQLEWIGRIYPOGDNNGYNGRFK 61
QY 66 DKATLTVKSSSTAYMELRSLTSEDSAVYICAR--STMITNTYMDYWGOGTSVTYSSAKT 123
DB 62 VRATLTAERSSTTVVHLSSLTSVDSAVYFCARFHYDYRRSYAMDYWGOGTSVTYSSAKT 121
QY 124 TTPSVYPLAPGSAQAOTNSHWTLGCLVKGYPPEPVTWNNGSLSSGCVHTFPVAVLOSPLYT 183
DB 122 TTPSVYPLAPGSAQAOTNSHWTLGCLVKGYPPEPVTWNNGSLSSGCVHTFPVAVLOSPLYT 181
QY 184 LSSSVTPSPSTWPSSETVTCNVAHPASSTKVDKKIIVPRD 221
DB 182 LSSSVTPSPSTWPSSETVTCNVAHPASSTKVDKKIIVPRD 219

RESULT 34
US-08-896-933-23
; Sequence 23, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896,933
; CURRENT FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 232

; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-23
Query Match 26.4%; Score 931.5; DB 4; Length 232;
Best Local Similarity 76.0%; Pred. No. 3.4e-55;
Matches 177; Conservative 21; Mismatches 34; Indels 1; Gaps 1;
QY 226 SEKSEINEKDLRKKSELOGTALGNLKOIYYNNKAITSSSEKSAQOFLTNTLLFKGFFTG 285
DB 1 SEKSEINEKDLRKKSELOGTALGNLKOIYYNNKAITSSSEKSAQOFLTNTLLFKGFFTG 59
QY 286 HPWYNDLLVDLGSATSEYEGSSVDLYGAYYGOCAGGTPNKTCMYGGVTLHNNRLT 345
DB 60 HSWYNDLLVDFDSKDIDVKYKGGKVDLYGAYYGOCAGGTPNKTCMYGGVTLHNNRLT 119
QY 346 EEKVPINLWIDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQ 405
DB 120 EEKVPINLWIDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQ 179
QY 406 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRONTTITSSLSLSLYLTT 458
DB 180 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRONTTITSSLSLSLYLTT 232

RESULT 35
US-09-314-235-23
; Sequence 23, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; CURRENT FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416,530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-23

Query Match 26.4%; Score 931.5; DB 4; Length 232;
Best Local Similarity 76.0%; Pred. No. 3.4e-55;
Matches 177; Conservative 21; Mismatches 34; Indels 1; Gaps 1;
QY 226 SEKSEINEKDLRKKSELOGTALGNLKOIYYNNKAITSSSEKSAQOFLTNTLLFKGFFTG 285
DB 1 SEKSEINEKDLRKKSELOGTALGNLKOIYYNNKAITSSSEKSAQOFLTNTLLFKGFFTG 59
QY 286 HPWYNDLLVDLGSATSEYEGSSVDLYGAYYGOCAGGTPNKTCMYGGVTLHNNRLT 345
DB 60 HSWYNDLLVDFDSKDIDVKYKGGKVDLYGAYYGOCAGGTPNKTCMYGGVTLHNNRLT 119
QY 346 EEKVPINLWIDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQ 405
DB 120 EEKVPINLWIDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQ 179
QY 406 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRONTTITSSLSLSLYLTT 458

Db 180 RGLIVFHTSTEPSVNDLFGAQQYNTLLRIYRDNKTINSENHDIYLYTS 232

RESULT 36

US-09-144-776B-2
; Sequence 2, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
; Mark A. Olson
; Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
; Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army MRC -504 Scott Street
; MCMR-JA (Charles H. Harris-Patent
; Atty)
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,776B
; FILING DATE: 01-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/882,431
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-144-776B-2

Query Match 26.3%; Score 925; DB 4; Length 257;
Best Local Similarity 75.1%; Pred. No. 1.le-54;
Matches 175; Conservative 21; Mismatches 37; Indels 0; Gaps 0;
QY 226 SEKSEINEKDLRKKSELOGTALGNLQIYYNKAITSSEKSADQFLNTLLFKGFFTG 285
Db 25 SEKSEINEKDLRKKSELOGTALGNLQIYYNKAITSSEKSADQFLNTLLFKGFFTG 84
QY 286 HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
Db 85 HSWYNDLLVRFDSKDIDYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKVPINLWIDGKQTVPIDKVTKSKKVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 405
Db 145 EEKVPINLWIDGKQTVPIDKVTKSKKVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 204
QY 406 RGLIVFHSSEGSTVSVDLFDAGQYPTDILLRIYRDNKTINSENHDIYLYTT 458
Db 205 RGLIVFHTSTEPSVNDLFGAQQYNTLLRIYRDNKTINSENHDIYLYTS 257

RESULT 37

US-09-144-776B-4
; Sequence 4, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
; Mark A. Olson
; Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
; Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army MRC -504 Scott Street
; MCMR-JA (Charles H. Harris-Patent
; Atty)
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,776B
; FILING DATE: 01-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/882,431
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-144-776B-4

Query Match 26.1%; Score 921; DB 4; Length 233;
Best Local Similarity 75.0%; Pred. No. 1.7e-54;
Matches 174; Conservative 21; Mismatches 37; Indels 0; Gaps 0;
QY 227 EKSEINEKDLRKKSELOGTALGNLQIYYNKAITSSEKSADQFLNTLLFKGFFTG 286
Db 2 EKSEINEKDLRKKSELOGTALGNLQIYYNKAITSSEKSADQFLNTLLFKGFFTG 61
QY 287 PWNLDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 346
Db 62 SWYNDLLVRFDSKDIDYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
QY 347 EEKVPINLWIDGKQTVPIDKVTKSKKVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 406
Db 122 EEKVPINLWIDGKQTVPIDKVTKSKKVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 181
QY 407 GLIVFHSSEGSTVSVDLFDAGQYPTDILLRIYRDNKTINSENHDIYLYTT 458
Db 182 GLIVFHTSTEPSVNDLFGAQQYNTLLRIYRDNKTINSENHDIYLYTS 233
RESULT 38
US-08-792-824-10

; Sequence 10, Application US/08792824
; Patent No. 5932449
; GENERAL INFORMATION:
; APPLICANT: EMANUEL, PETER A.
; APPLICANT: BURANS, JAMES P.
; APPLICANT: VALDES, JAMES J.
; APPLICANT: MOHVEE, ELDEFRAWI E.
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. Army Chemical and Biological Defense
; STREET: Office of the Chief Counsel, Bldg E4435
; CITY: Aberdeen Proving Ground
; STATE: MD
; COUNTRY: U.S.
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08792.824
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Biffoni, U. J.
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 431-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158
; TELEFAX: 410-671-2534
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-792-824-10

Query Match 26.0%; Score 915.5; DB 2; Length 254;
Best Local Similarity 79.2%; Pred. No. 4.5e-54;
Matches 179; Conservative 15; Mismatches 27; Indels 5; Gaps 2;
QY 1 EVLOQSGPDLVKPGASVKISKASGYSTGYMHVWVKQSPKGLGWIGRINPNNGVTLY 60
DB 23 EVLOQSGAELVKPGASVKLSCTASGFNIKDTFMMHWKQRPQEGLEWIGRIDPANGNTEY 82
QY 61 NOKFKDKATLTVDKSSTTAYMELRSLTSEDSAVYYCARSTMTNVMYDVGOGTSVTYSS 120
DB 83 DPKFGKATITADTSSNTVNLQSLTSEDYAVYYCASGGELG---FPYWGQGLTLTVSA 139
QY 121 AKTTSPSVYPLAPGSAQAQNSMVTGLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
DB 140 AKTTSPSVYPLAPGSAQAQNSMVTGLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 199
QY 181 LYTSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIIPRD---SGG 224
DB 200 LYTSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIIPRDCTSGG 245

RESULT 39
US-08-792-824-13
; Sequence 13, Application US/08792824
; Patent No. 5932449
; GENERAL INFORMATION:
; APPLICANT: EMANUEL, PETER A.
; APPLICANT: BURANS, JAMES P.
; APPLICANT: VALDES, JAMES J.
; APPLICANT: MOHVEE, ELDEFRAWI E.
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN
; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. Army Chemical and Biological Defense
; STREET: Office of the Chief Counsel, Bldg E4435
; CITY: Aberdeen Proving Ground
; STATE: MD
; COUNTRY: U.S.
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08792.824
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Biffoni, U. J.
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 431-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158
; TELEFAX: 410-671-2534
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-792-824-13

Query Match 26.0%; Score 915.5; DB 2; Length 254;
Best Local Similarity 79.2%; Pred. No. 4.5e-54;
Matches 179; Conservative 15; Mismatches 27; Indels 5; Gaps 2;
QY 1 EVLOQSGPDLVKPGASVKISKASGYSTGYMHVWVKQSPKGLGWIGRINPNNGVTLY 60
DB 23 EVLOQSGAELVKPGASVKLSCTASGFNIKDTFMMHWKQRPQEGLEWIGRIDPANGNTEY 82
QY 61 NOKFKDKATLTVDKSSTTAYMELRSLTSEDSAVYYCARSTMTNVMYDVGOGTSVTYSS 120
DB 83 DPKFGKATITADTSSNTVNLQSLTSEDYAVYYCASGGELG---FPYWGQGLTLTVSA 139
QY 121 AKTTSPSVYPLAPGSAQAQNSMVTGLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
DB 140 AKTTSPSVYPLAPGSAQAQNSMVTGLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 199
QY 181 LYTSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIIPRD---SGG 224
DB 200 LYTSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIIPRDCTSGG 245

RESULT 40
US-08-792-824-4
; Sequence 4, Application US/08792824
; Patent No. 5932449
; GENERAL INFORMATION:
; APPLICANT: EMANUEL, PETER A.
; APPLICANT: BURANS, JAMES P.
; APPLICANT: VALDES, JAMES J.
; APPLICANT: MOHVEE, ELDEFRAWI E.
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. Army Chemical and Biological Defense
; STREET: Office of the Chief Counsel, Bldg E4435
; CITY: Aberdeen Proving Ground
; STATE: MD
; COUNTRY: U.S.
; ZIP: 21010-5423
; COMPUTER READABLE FORM:

Db 138 TVSIFPPSSBQLTSGGASVWCFLNNFYPKDINVKKIDGSRQNGVLNSWTDQDQSKDSTY 197

Qy 632 SMSSTLTLTCKDEYERHNSYTCEATHKTSPIVKSFNRE 671

Db 198 SMSSTLTLTCKDEYERHNSYTCEATHKTSPIVKSFNRE 237

RESULT 43

5455030-1

Patent No. 5455030

APPLICANT: LADNER, ROBERT C.; BIRD, ROBERT E.; HARDMAN, KARL

TITLE OF INVENTION: IMMUNOTHERAPY USING SINGLE CHAIN

POLYPEPTIDE BINDING MOLECULES

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/40,440

FILING DATE: 1-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 512,910

FILING DATE: 25-APR-1990

APPLICATION NUMBER: 299,617

FILING DATE: 19-JAN-1989

APPLICATION NUMBER: 92,110

FILING DATE: 02-SEP-1987

APPLICATION NUMBER: 902,971

FILING DATE: 01-SEP-1986

SEQ ID NO:1

LENGTH: 447

5455030-1

Query Match 25.7%; Score 903.5; DB 6; Length 447;

Best Local Similarity 32.9%; Pred. No. 5.9e-53;

Matches 222; Conservative 69; Mismatches 121; Indels 263; Gaps 15;

Qy 1 EVLOQSGPDLVKPGASVKISKASGYSFTGYHHWYKQSPGKLGLEWIGRINPNNGVTLY 60

Db 1 EVLOVSGDGLVKPG-SLKLASAGSTFTISYGMWVRQTPDKRLEWVATISSGSTYTY 59

Qy 61 NQKFKDRLATLVKDSSTTAYMELASLASEDSAVYICAR---STMTN-YVMDYWGQGTSV 116

Db 60 PDSVKGRTISRDNKNTLYLQMSGLKSEDTAMYICARRITTVLTVDYANDYWGQGTSV 119

Qy 117 TVSSAKTTPPSVYPLAPGSAQAQNSMVTLGCLVKGYPPEPVTVTNWNSGLSGVHTFPAP 176

Db 120 TVSSAKTTPPSVYPLAPGSAQAQNSMVTLGCLVKGYPPEPVTVTNWNSGLSGVHTFPAP 179

Qy 177 LQSDLYTLSSSVTPSSWPSETVTCNVAHPASSTKVDKIVPRDSGSPSEKSEINEKD 236

Db 180 LQSDLYTLSSSVTPSSWPSETVTCNVAHPASSTKVDKIVPRDCG----- 226

Qy 237 LRKKESELQGTALGNLKQIYYNNSKAITSSSEKSADQFLTNTLLFKGFTGHPWYNDLLVDL 296

Db 227 ---CKPCICIVPEVSSVFIIPPK-----PKDVLITL----- 255

Qy 297 GSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMGVGVTLHDNRLTEKKYPINLWI 356

Db 256 -----TP-KVTCVVVDIS-----KDDPEVQFSWFV 279

Qy 357 DGKQTTVPIDKVKTSKKEV-----TVQELDLQARHYLHCKFGLYNSDS--FGKKVORG 407

Db 280 D-----DVEHTAQTPREEQDTSRSVPELPIHQDMLNKEFKCRVDVSAAFPAPIEK- 334

Qy 408 LIVFHSSEGSTSVSYDLFDAQGOYPTDLLRIYRDNNTTISSTLSISLYLYTTSIVMTQPT 467

Db 335 -----TISKT----- 339

Qy 468 SLLVSAGDRVTTCKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGT 527

Db 340 ----- 339

Qy 528 DFTLTISVQAEADAAYFCQDYNPPTFGGCTKLEIKRADAAPTYSIFPPSSQLTSGG 587

Db 340 -----KGRKAPQVYTIIPPPEQMAKDK 362

Qy 588 ASVVCFLNNFYPKDINVKKIDGSRQNGVLNSWTDQDQSKDSTYSMSSTLTLTCKD 642

Db 363 VSLTCMITDFPEDITVEMQWDGQPAENYKNTQRIIMNT-----DGSYFVSKLDVQKS 415

Qy 643 EYERHNSYTCEATHK 657

Db 416 NWEAGDTFTCSVLHE 430

RESULT 44

PCT-US94-14106-57

Sequence 57, Application PC/TUS9414106

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Process for Generating Specific Antibodies

NUMBER OF SEQUENCES: 61

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII (text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/14106

FILING DATE:

CLASSIFICATION:

INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:

LENGTH: 218 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US94-14106-57

Query Match 25.6%; Score 901; DB 5; Length 218;

Best Local Similarity 80.6%; Pred. No. 3.5e-53;

Matches 174; Conservative 9; Mismatches 29; Indels 4; Gaps 1;

Qy 460 IVMTQTPTSLVLSAGDRVITITCKASQSVSND---VAVYQKPGQSPKLLISYTSRVRAG 515

Db 2 IVMTQSPASLAVSLGQRATISCRVRQSVSTSSHVMHWYQKPGQPPKLLIKYASNLESG 61

Qy 516 VPDFSGSGYGTDTFTLTISVQAEADAAYFCQDYNPPTFGGCTKLEIKRADAAPTYSI 575

Db 62 VPAREFSGSGGTDTFTLTINHPVEEEDATYYCQHSWEIPTYETGGGTCKLEIKRADAAPTYSI 121

Qy 576 FPPSEQLTSGGASVVCFLNNFYPKDINVKKIDGSRQNGVLNSWTDQDQSKDSTYSMS 635

Db 122 FPPSEQLTSGGASVVCFLNNFYPKDINVKKIDGSRQNGVLNSWTDQDQSKDSTYSRS 181

Qy 636 TLTLTCKDEYERHNSYTCEATHKTSPIVKSFNRE 671

Db 182 TLTLTCKDEYERHNSYTCEATHKTSPIVKSFNRE 217

RESULT 45

US-09-170-769A-8

Sequence 8, Application US/09170769A

Patent No. 6444206

GENERAL INFORMATION:

APPLICANT: THE SCRIPPS RESEARCH INSTITUTE

APPLICANT: LETURCO, Didier

APPLICANT: MORIATRY, Ann

APPLICANT: ULEVITCH, Richard

APPLICANT: TOBIAS, Peter

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING CD14 MEDIATED CELL ACT

FILE REFERENCE: SCRIP1140-3

CURRENT APPLICATION NUMBER: US/09/170,769A

CURRENT FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: US 08/070,160

PRIOR FILING DATE: 1993-05-28

NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn version 3.0

```
; SEQ ID NO 8
; LENGTH: 211
; TYPE: prt
; ORGANISM: Murine
US-09-170-769A-8

Query Match          25.5%; Score 898; DB 4; Length 211;
Best Local Similarity 81.4%; Pred. No. 5.3e-53;
Matches 171; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

Qy 462 MTQPTSLIVSAGDRVTITCKASQSVNDVAVYQKQSPKLLISYTSRYAGVPDRS 521
Db 1 MTQPTSSLSASLGRVITISCRASQDKNYLMWYQGGPGTVKVLIIYTSRLHSGVPSR 60
Qy 522 GSGYGTDTLTITSSVOAEDAAYVFCQDYNPSPTFGGKLEIKRADAAAPTIVSIFPPS 581
Db 61 GSGGTDYSLTISNLEQDFATYFCQRGDTLPWTFGGTKLEIKTADAAAPTIVSIFPPS 120
Qy 582 QLTSGGASVVCFLNNFYPKDINVKKIDGSRQGVLSNWTQDQSKDSTYSMSSTLT 641
Db 121 QLTSGGASVVCFLNNFYPKDINVKKIDGSRQGVLSNWTQDQSKDSTYSMSSTLT 180
Qy 642 DEYERHNSYTCETHKTSPIVKSFRNE 671
Db 181 DEYERHNSYTCETHKTSPIVKSFRNE 210

RESULT 46
PCT-US94-14106-55
; Sequence 55, Application PC/TUS9414106
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Process for Generating Specific Antibodies
; NUMBER OF SEQUENCES: 61
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14106
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-14106-55

Query Match          25.5%; Score 897.5; DB 5; Length 223;
Best Local Similarity 79.4%; Pred. No. 6.1e-53;
Matches 177; Conservative 12; Mismatches 31; Indels 3; Gaps 2;

Qy 1 EVLQQSGPDLVVKPGASVKISKAGSYSTGYMHWKQSPGKGLWGIRNPNGVTLY 60
Db 1 EVLQQSGAELMMPGASVKISKATGYTLSSWLEWYKQSPGCHLEWLGEILFGSGSAHY 60
Qy 61 NQKFKDQATLVDRKSSSTAYMELRSLTSEDSAVYYCARSTMITNY--YMDYWGQGTSTV 118
Db 61 NEKFKGKATFTVDTSSNTAYMQLSSLTSEDSAVYYCARGD--YGNVGDYFDYWGQGLTV 119
Qy 119 SSATKTPPSVYPLAGSAAQNTSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFPV 178
Db 120 FSAKTPFSSVYPLAGSAAQNTSMVTGLCLVKGYLPEVPTVTVTWNSGSLSSGVHFP 179
Qy 179 SDLYTLSSSVTPSPSTPSETVTCNVAHPASSTKVDRKIVPRD 221
Db 180 SDLYTLRSVTPSPSTPSETVTCNVAHPASSTKVDRKIVPRD 222

RESULT 47
```

```
US-08-737-129A-8
; Sequence 8, Application US/08737129A
; Patent No. 5885816
; GENERAL INFORMATION:
; APPLICANT: Ikuo FUJII et al.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES ENANTIOSELECTIVELY
; HYDROLYSING AMINO ACID ESTER DERIVATIVES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,129A
; FILING DATE: No. 5885816ember 15, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-737-129A-8

Query Match          25.4%; Score 895; DB 2; Length 215;
Best Local Similarity 80.2%; Pred. No. 8.5e-53;
Matches 170; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

Qy 460 IVMTQPTSLIVSAGDRVTITCKASQSVNDVAVYQKQSPKLLISYTSRYAGVPDR 519
Db 2 LVMTQPTSSMYASLGERVTITCKASQDINIYLSWFQOKPGKSPKALIVRTNGLVDGVPSR 61
Qy 520 FSGSGYGTDTLTITSSVOAEDAAYVFCQDYNPSPTFGGKLEIKRADAAAPTIVSIFPPS 579
Db 62 FSGSGSQDYSLTISSEYEDMWGIYVCIQYDFEYVTFGGTGKLEIKRADAAAPTIVSIFPPS 121
Qy 580 SEQLTSGASVVCFLNNFYPKDINVKKIDGSRQGVLSNWTQDQSKDSTYSMSSTLT 639
Db 122 SEQLTSGASVVCFLNNFYPKDINVKKIDGSRQGVLSNWTQDQSKDSTYSMSSTLT 181
Qy 640 TKDEYERHNSYTCETHKTSPIVKSFRNE 671
Db 182 TKDEYERHNSYTCETHKTSPIVKSFRNE 213

RESULT 48
US-08-442-542-18
; Sequence 18, Application US/08442542
; Patent No. 5685600
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine B.
; APPLICANT: Kozziel, Michael G.
; TITLE OF INVENTION: Antibodies which Bind to Insect Gut
```


Db 272 SSKATTPSVYPLAPGSRSAQAQNSMVTLCGLVKGYPFPTVTWNSGLSSGVHTFPV 331
Qy 177 LOSDLATLSSSVTPSPSTWPTETVCNVAHPASSTKVDKIVPRDSGCPSEKSEEINEKD 236
Db 332 LOSDLATLSSSVTPSPSTWPTETVCNVAHPASSTKVDKIVPRDCG----- 378
Qy 237 LRKKELOGTLGNLQIYYNSKAITSSSEKSADQFTLTLFKGFFTGHPWYNDLLVDL 296
Db 379 ---CKPCICIVPEVSSVFIEPPK-----PKDVLITL----- 407
Qy 297 GSTAATSEYSSVDLYGAYGYOCAGTGNKTCAGMYGGVTLHDNRLTEKKYVINLWI 356
Db 408 -----TP-KVTCVVVDIS-----KDDPEVQFSWFV 431
Qy 357 DGKQTTVPIDKVKTSKEV-----TVQELDLQARHVLHGK-FGL-YNSDSFGKVGORG 407
Db 432 D-----DVEHTAQTQPREEQNFSTRSELPIMHQDLNGKEFKCRVNSAAFPAPIEK- 486
Qy 408 LIVFHSSEGSTVSYDLFDAQGYPTLLRIYRDNTTISSTSLISISLYLYTTTSIVMTQPT 467
Db 487 -----TISK----- 491
Qy 468 SLLVSAGDRVTTCKASQSVNDVAMVQKPGQSPKLLISYTSRYAGVPRFSGSGYGT 527
Db 492 ----- 491
Qy 528 DFTLTSSVQAEADAAVFCQODYNSPPTFGGKLEIKRADAAAPTYSIFPPSSQOLTS 587
Db 492 -----KGRPKAPQVYTIPTPPKEQMAKDK 514
Qy 588 ASVVCFLNNFYPKIDINVKWKIDGSRQ-----GVLSNWTQDQSKDSTYSMSSTLTITKD 642
Db 515 VSLTCMTIDFPEDITVEWQNGQPAENYKNTQPMNT-----NGSYFYVSKLNVQKS 567
Qy 643 EYERHNSYTCEATHK 657
Db 568 NWEAGNTEFTCSVLHE 582

RESULT 50

US-09-423-439-58
; Sequence 58, Application US/09423439
; Patent No. 6339070
; GENERAL INFORMATION:
; APPLICANT: EMERY, Stephen Charles
; TITLE OF INVENTION: CHEMICAL COMPOUNDS
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Winthrop, L.L.P.
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423.439
; FILING DATE: 09-No. 6339070-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01294
; FILING DATE: 05-MAY-1998
; APPLICATION NUMBER: GB 9709421.3
; FILING DATE: 10-MAY-1997
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-09-423-439-58

Query Match 25.3%; Score 889.5; DB 4; Length 235;
Best Local Similarity 77.1%; Pred. No. 2.2e-52;
Matches 172; Conservative 19; Mismatches 31; Indels 1; Gaps 1;

Qy 449 LSISLYLYTTSIVMTQTPTSLLSVAGDRVTTITCKASQSVNDVAMVQKPGQSPKLLISY 508
Db 13 ISASVIMSRGQTVLSQSPAILSASPGKVTMTCRASSVTY-IHWYQOKPGSSPKSWIYA 71
Qy 509 TSSRYAGVPRDFSGSGYCTDFTLTSSVQAEADAAVFCQODYNSPPTFGGKLEIKRAD 568
Db 72 TSNLASGVPARFSGSGSGTYSLTISRVEADAATYYCQHWSSSKPPTFGGKLEIKRAD 131
Qy 569 AAPTYSIFPPSSQOLTSQGSASVVCFLNNFYPKIDINVKWKIDGSRQNGVLNSWTDQDSKD 628
Db 132 AAPTYSIFPPSSQOLTSQGSASVVCFLNNFYPKIDINVKWKIDGSRQNGVLNSWTDQDSKD 191
Qy 629 STYSMSSTLTITKDEYERHNSYTCEATHKSTSPIVKSFNRNE 671
Db 192 STYSMSSTLTITKDEYERHNSYTCEATHKSTSPIVKSFNRNE 234

Search completed: January 8, 2003, 11:59:34
Job time : 30.0486 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:58:32 ; Search time 12.6232 Seconds
(without alignments)
1032.821 Million cell updates/sec

Title: US-09-900-766-1

Perfect score: 3522

Sequence: 1 EVQLQQSGPDLVKPGASVKI.....EATHKTSPIVKSNRNES 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1257	35.7	669	US-09-807-721-2	Sequence 2, Appl
2	1072	30.4	248	US-09-870-759-16	Sequence 16, Appl
3	1028	29.2	597	US-09-903-327A-11	Sequence 11, Appl
4	1024.5	29.1	613	US-09-903-327A-14	Sequence 14, Appl
5	1023.5	29.1	456	US-09-903-327A-2	Sequence 2, Appl
6	1023.5	29.1	493	US-09-903-327A-13	Sequence 13, Appl
7	1023.5	29.1	510	US-09-903-327A-12	Sequence 12, Appl
8	1022.5	29.0	438	US-09-903-327A-6	Sequence 6, Appl
9	935	26.5	257	US-09-870-759-8	Sequence 8, Appl
10	912	25.9	238	US-09-903-327A-4	Sequence 4, Appl
11	855.5	24.3	473	US-09-910-059-131	Sequence 131, Appl
12	807	22.9	448	US-09-917-410-6	Sequence 6, Appl
13	782	22.5	461	US-09-249-011A-24	Sequence 24, Appl
14	789	22.4	473	US-09-828-995B-20	Sequence 20, Appl
15	789	22.4	729	US-09-825-012-52	Sequence 52, Appl
16	789	22.4	730	US-09-825-012-49	Sequence 49, Appl
17	789	22.4	731	US-09-825-012-46	Sequence 46, Appl
18	789	22.4	739	US-09-825-012-61	Sequence 61, Appl
19	789	22.4	740	US-09-825-012-58	Sequence 58, Appl

20	789	22.4	741	10	US-09-825-012-55	Sequence 55, Appl
21	785.5	22.3	470	10	US-09-859-053-28	Sequence 28, Appl
22	780	22.1	230	10	US-09-995-693-2	Sequence 2, Appl
23	779.5	22.1	476	10	US-09-747-669-3	Sequence 3, Appl
24	776.5	22.0	232	10	US-09-754-998-2	Sequence 2, Appl
25	776.5	22.0	468	9	US-09-828-995B-5	Sequence 5, Appl
26	759.5	21.6	470	9	US-09-828-995B-11	Sequence 11, Appl
27	756	21.5	450	10	US-09-811-384-12	Sequence 12, Appl
28	740	21.0	240	10	US-09-799-514-8	Sequence 8, Appl
29	728	20.7	237	10	US-09-940-166A-6	Sequence 6, Appl
30	726	20.6	214	10	US-09-940-166A-2	Sequence 2, Appl
31	726	20.6	214	10	US-09-811-384-11	Sequence 11, Appl
32	726	20.6	220	10	US-09-995-693-1	Sequence 1, Appl
33	725.5	20.6	300	10	US-09-940-166A-7	Sequence 7, Appl
34	722.5	20.5	515	10	US-09-825-012-66	Sequence 66, Appl
35	722.5	20.5	517	10	US-09-825-012-38	Sequence 38, Appl
36	722.5	20.5	519	10	US-09-825-012-76	Sequence 76, Appl
37	722.5	20.5	519	10	US-09-825-012-80	Sequence 80, Appl
38	722.5	20.5	521	10	US-09-825-012-71	Sequence 71, Appl
39	722.5	20.5	525	10	US-09-825-012-85	Sequence 85, Appl
40	722.5	20.5	527	10	US-09-825-012-43	Sequence 43, Appl
41	722.5	20.5	529	10	US-09-825-012-95	Sequence 95, Appl
42	722.5	20.5	531	10	US-09-825-012-90	Sequence 90, Appl
43	718.5	20.4	232	10	US-09-811-384-10	Sequence 10, Appl
44	718.5	20.4	241	10	US-09-940-166A-1	Sequence 1, Appl
45	718.5	20.4	449	10	US-09-736-371B-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-09-807-721-2

: Sequence 2, Application US/09807721

: Patent No. US20020174453A1

: GENERAL INFORMATION:

: APPLICANT: AUBURN UNIVERSITY

: APPLICANT: UNIVERSITY OF CENTRAL FLORIDA

: TITLE OF INVENTION: PRODUCTION OF ANTIBODIES IN TRANSGENIC PLASTIDS

: FILE REFERENCE: 1463-PCT-US-00

: CURRENT APPLICATION NUMBER: US/09/807,721

: CURRENT FILING DATE: 2001-12-21

: PRIOR APPLICATION NUMBER: PCT/US01/06274

: PRIOR FILING DATE: 2001-02-28

: NUMBER OF SEQ ID NOS: 6

: SOFTWARE: PatentIn Ver. 2.1

: SEQ ID NO 2

: TYPE: PRT

: ORGANISM: Homo sapiens

: US-09-807-721-2

Query Match 35.7%; Score 1257; DB 9; Length 669;
Best Local Similarity 41.6%; Pred. No. 3.6e-47;
Matches 308; Conservative 87; Mismatches 201; Indels 144; Gaps 23;

QY 2 VOLQQSGPDLVKPGASVKISCRASGYFTGYMHVWQSPGKLEWIGRINPNNGVTLYN 61

Db 3 VOLQQSGPDLVKPGASVKISCRASGYFTFDYNIHWVQSRGKSLWIGIYYPYNGTYN 62

QY 62 QRFKDKATITVDKSSITAYMELRSLTSEDSAVVYCARSTMTNVMYMGQGTSTVTSSA 121

Db 63 QRFKDKATITVDNSSTTSAYMELRSLTSEDSAVVYCAT-----YFDYMGQGTTLTVSSA 115

QY 122 KITPPSVYPLAFGSAATSNMTGLCLVKGYFP-EPVTVWNSGSLSSGVHFFPAV--LQ 178

Db 116 SPTSPKVFPLSLDSTPDQGNV--VACLVGQFFPEPLSVTWSGQNVARNFPSPQDAS 174

QY 179 SDLTTLSSSVTPSPSSWNP-SETVTGNVAHPASSTKVDDKIVPRDSGGPSEKSEINEKDL 237

Db 175 GDLYTTSSQLTLPATCPDGKSVTCHVKHYTNSSQ--DVTVPGRVPPPP-----PCC 224

QY 238 RKKSELQGTALGNLKOIYYYNKAITTSEKSAQDLTNTLLFKGFTTGHVWYNDLLVDLG 297

Db 225 HPLSLHRALEDL-----LIGSEAN-----LTCTL-----TG-----LRDASG 258
QY 298 STAATSEYEGSSV-----DLYGAY-----YGCAGGTGPNKTACMYG 334
Db 259 ATFTWTFSPGKSAVOPPERDLGCGYSVSVLPGCAQPNWHGETFTCTAAHPELKTPLTA 318
QY 335 GVTLDHNNRLTEKKVP-----INLWIDGKQTTVPIDKV 368
Db 319 NITKSGNTFRPEVHLPLPSEELALNELVTLCLARGFSPKDVLRWLGSGQ-ELPREKY 377
QY 369 KT--SKKE-----VTQELDLQARHYLHGKFGLYNSDFGGKQVORGLIVFHSSEGST 418
Db 378 LTWASRQESQGTITTVAVTSILRVAEDHKKG-----ETP-----SCMVGHEALPLA 424
QY 419 VSVDLFDAOGQYPTDILLRIYRONTTISSTLSISLYL-----YPTSVMTOTPTSLLV 471
Db 425 FTQKTTIDRLAGKP-----THINVSVMAEADGTCYRMDIVMTQSPAIMSA 469
QY 472 SAGDRVTITCKASQSVSNDVANYQKPGOSPKLLISYTSRAGYVDPDRFSGSGYGTDTTL 531
Db 470 SPGEKVTITCSASSMWVSY-MHMFQOKPGTSPKLLWLYSTNLASGVPAREFSGSGSTSYSL 528
QY 532 TISSVOAEDAAYVFCQDYNSTPFGGGTKLEIKRADAAPTVSIFFPSSEQLTSGGASVV 591
Db 529 TISRMEADAATYVCHORTSYPTFGGGTKLEIKRTVAAPSVFIFFPSDEQLKSGTASVV 588
QY 592 CFLNNYPKDNVKKIDGSRONGVLNSWTQDSDKDYSSMSLTLLTKDEYERHNSYT 651
Db 589 CLLNNFYPREAKVQWKNVNALQSGNSQESVTEQDSKDYSSLSSTLTLTKADYEHKHYA 648
QY 652 CEATHKTSPTPIKSPNRNE 671
Db 649 CEVTHGLSSPVTKSPNRGE 668

RESULT 2
US-09-870-759-16
; Sequence 16, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-870-759-16

Query Match 30.4%; Score 1072; DB 9; Length 248;
Best Local Similarity 90.2%; Pred. No. 1.le-39;
Matches 202; Conservative 9; Mismatches 13; Indels 0; Gaps 0;
QY 226 SEKSEINEKDLRKKSELOGTALGNLKOIYYNNSKAITSEKSADQFLNTLLFKGFFTG 285
Db 25 SEKSEINEKDLRKKSELORNALNLRQIYYNEKAITENKESDDQFLNTLLFKGFFTG 84
QY 286 HPWYNLLVDLGSTAATSEYSGSSVDLYGAYGYQCAGGTGPNKTACMYGGVTLHDNNRLT 345
Db 85 HPWYNLLVDLGSKDATNRYKGGKVDLYGAYGYQCAGGTGPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKKVPINLWIDGKQTTVPIDKVTSSKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQ 405
Db 145 EEKKVPINLWIDGKQTTVPIDKVTSSKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 406 RGLIVFHSSEGSTVSYDLFDAOGQYPTDILLRIYRONTTISSTLSL 449

Db 205 RGLIVFHSSEGSTVSYDLFDAOGQYPTDILLRIYRONTKNTINSEN 248
RESULT 3
US-09-903-327A-11
; Sequence 11, Application US/09903327A
; Patent No. US2002016433A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erquang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TAR
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and TNF alpha mature peptide
US-09-903-327A-11

Query Match 29.2%; Score 1028; DB 9; Length 597;
Best Local Similarity 37.2%; Pred. No. 1.7e-37;
Matches 253; Conservative 53; Mismatches 115; Indels 260; Gaps 15;
QY 1 EVQLQQSGPDLYKPGASVKISCKASGYSTGYIMHWKQSPGKLEWIGRINPNNGVTLY 60
Db 20 EVQLQQSGPELYKPGASVKISCKASGYSTGYIMHWKQSHGKSLWIGIYIYKGGTGY 79
QY 61 NOKFKDKATLVKDSSTTAYMELRSLTSDSAVYYCARSTMTNTNYMDYWGOGTSTVYSS 120
Db 80 NOKFKSKATLTDTSSNTAYMELRSLTSDSAVYYCARG-----IAYWGGGTTLTVSA 132
QY 121 AKTTPSPVYPLAPGSAAGTNSMTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPFVAVLQSD 180
Db 133 AKTTPSPVYPLAPGSAAGTNSMTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPFVAVLQSD 192
QY 181 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIIVPRDSGGPSEKSEINEKDLRKK 240
Db 193 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIIVPRDCG-----235
QY 241 SELOGTALGNLKOIYYNNSKAITSEKSADQFLNTLLFKGFFTGHPWYNLLVDLGSTA 300
Db 236 CKPCICTVPEVSSVFIFPPK-----PKDVLITL-----264
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTGPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
Db 265 -----TP-KVTCVVVDIS-----KDDPEVQSFWEVD-----289
QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKQVQRLIVF 411
Db 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHODLWLNKREKFCRVNSAAPPAPIEK-----343
QY 412 HSSEGSTVSYDLFDAOGQYPTDILLRIYRONTTISSTLSISLYLTYTTSIVMTQTPTSLLV 471
Db 344 -----TISK-----348
QY 472 SAGDRVTITCKASQSVSNDVANYQKPGOSPKLLISYTSRAGYVDPDRFSGSGYGTDTFL 531
Db 349 -----348
QY 532 TISSVOAEDAAYVFCQDYNSTPFGGGTKLEIKRADAAPTVSIFFPSSEQLTSGGASVV 591
Db 349 -----KGRAPAQVYTTIPPPKEQMAKDKVSLT 375

Qy	592	CFLNFPKDKINVKKIDGSEKQGVLSWTDDSKDSYKSSSTLTLLKDEYERINST	651
Db	376	CMITDFPDDITVEWQNGQPAEN-YKNTQPIMDT-DGSFYVSKLVKSNKNEAGNTEI	433
Qy	652	CEATHKTSTSPVKSFNRNES	672
Db	434	CSVLHE-----FVRRSSSRTPS	449

```

RESULT 4
US-09-903-327A-14
; Sequence 14, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erguang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and SCF mature peptide
US-09-903-327A-14

```

Query Match 29.1%; Score 1024.5; DB 9; Length 613;
Best Local Similarity 37.3%; Pred. No. 2.5e-37;
Matches 250; Conservative 51; Mismatches 115; Indels 255; Gaps 14;

Qy	1	EVOLQSGPDLVKPGASVKISCKASGYSTFYVYHHVVKSPGKGLIEWIRINPNNGVTLY	60
Db	20	EVOLQSGPELVLKPGASVKISCKASGYTFDYNMHWVKSHGKSLIEWIYIYPYKGGTGY	79
Qy	61	NOFKDKATLTVDKSSNTAYMELRSLTSEDSAVYYCARSTMITNYVMDYWGQGTSTVTS	120
Db	80	NOFKSKATLTDTSSNTAYMELRSLTSDASAVYYCARG-----IAYWGQGLTLYT	132
Qy	121	AKTTPSPVYPLPGSAAQNSMTVLCGLVKGYPEPEPVYTWNSGSLSSGVHFTPAVLQSD	180
Db	133	AKTTPSPVYPLPGSAAQNSMTVLCGLVKGYPEPEPVYTWNSGSLSSGVHFTPAVLQSD	192
Qy	181	LYTLSSSVTPSPSTWPSSETVTCNVAHPASSTKVDDKIVPRDSCGPPSEKSEETNEKDLRK	240
Db	193	LYTLSSSVTPSPSTWPSSETVTCNVAHPASSTKVDDKIVPRDCG-----	235
Qy	241	SELQGTALGNLKIYYNSKAITSSSEKSAQDFLTNTLLFKGFTTGHFWYNDLLDLGSTA	300
Db	236	CKPCICITVPEVSSVFIPPK-----PKDVLITL-----	264
Qy	301	ATSEYEGSSVDLYGAYVGYQCAGGTGNKTYACMYGGVTLHDNNRLTEEKVPINLWDGKQ	360
Db	265	-----TP-KVTCVVVDIS-----KQDPEVQPSWFVD---	289
Qy	361	TTVPIDKVTSSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSPGGKVQRGOLIV	411
Db	290	-DVEVHTAQTPREEQFNSTFRSVSELPIMHQDLWLNKGKFKRVNSAAFPAPIEK-----	343
Qy	412	HSSEGSTSVSYDLFDAQCGYPTDLLRIYRONTTTISSTLSLSLYLYTTSIVMTQTPTSLLV	471
Db	344	-----TISK-----	348
Qy	472	SAGDRVITICKASQSVSDNVAWYQKPGQSPKLLISYTSRIVAGVPDRFSGSGYGTDFTL	531

Db 290 -DVEVHTAQTPREEQFNSTFRSVSELPIMHQDWLNGKFKCRVNSAAPPAPIEK----- 343
QY 412 HSSEGSTVSYDLFDAQGOYPTDILLRIYRDNTTISLSISLYLTTISVMTQTPTSLLV 471
Db 344 -----TISK----- 348
QY 472 SAGDRVTITCASQSVNDVAVYQKPGQSKLLISYSSRYAGVDPDRFSGSGYGTDFTL 531
Db 349 ----- 348
QY 532 TISSVOAEDAAVYFCQDYNSTPTEGGGKLEIKRADAAPTYSIPPSSEQLTSGGASV 591
Db 349 -----KGRKPAQVYTIPTPPKEQMAKDVKSLT 375
QY 592 CFLNFPYKIDINVKWKIDGSRQNGVLSWTDQDSDKSTYSMSSTLTTLTKDEYERHNSYT 651
Db 376 CMITDFPEDITVEQWNGQPAEN-YKNTQPIMDT-DGSYFYVSKLVNOKSNWEAGNTFI 433
QY 652 CEATHK 657
Db 434 CSVLHE 439
RESULT 6
US-09-903-327A-13
; Sequence 13, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and EGF mature peptide
US-09-903-327A-13

Query Match 29.1%; Score 1023.5; DB 9; Length 493;
Best Local Similarity 37.4%; Pred. No. 2.3e-37;
Matches 249; Conservative 51; Mismatches 111; Indels 255; Gaps 14;
QY 1 EVQLQQSGDPLVKPGASVKISKASGYSTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db 20 EVQLQQSGPELVKPGASVKISKASGYFTDYNMHWKQSHGKSLWIGIYIPYKGGTG 79
QY 61 NQKFKDKATLTVDKSSRAYMELRLTSDSAVYYCARSTMITNTVMYWGOGTSTVYSS 120
Db 80 NQKFKSKATLTDDSSNTAYMELRLTSDSAVYYCARG-----IAYWGGTGLTVTVA 132
QY 121 AKTTPSPVPLAPGSAQAQNSMVTLCCLVKGYPPEPVTVTNWNSGSLSSGVHTFPVAVLQSD 180
Db 133 AKTTPSPVPLAPGSAQAQNSMVTLCCLVKGYPPEPVTVTNWNSGSLSSGVHTFPVAVLQSD 192
QY 181 LYTSSSVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDGGPSEKSEINEKDLRKK 240
Db 193 LYTSSSVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDGG----- 235
QY 241 SELOGTALGNLQIYYYSKAITSSSEKSAQDQFLTWTLFKGFTTCHPWYNDLLDLGSTA 300
Db 236 CRPCICTVPEVSVFIPPPK-----PKDVLTLT----- 264

QY 301 ATSEYEGSVLDLYGAYGYQCAGGTPNKACTMYGGVTLHDNNRLTEKKVPIINLWIDGKQ 360
Db 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD--- 289
QY 361 TTVPIDKYTKSKKEY-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVGKQLIVF 411
Db 290 -DVEVHTAQTPREEQFNSTFRSVSELPIMHQDWLNGKFKCRVNSAAPPAPIEK----- 343
QY 412 HSSEGSTVSYDLFDAQGOYPTDILLRIYRDNTTISLSISLYLTTISVMTQTPTSLLV 471
Db 344 -----TISK----- 348
QY 472 SAGDRVTITCASQSVNDVAVYQKPGQSKLLISYSSRYAGVDPDRFSGSGYGTDFTL 531
Db 349 ----- 348
QY 532 TISSVOAEDAAVYFCQDYNSTPTEGGGKLEIKRADAAPTYSIPPSSEQLTSGGASV 591
Db 349 -----KGRKPAQVYTIPTPPKEQMAKDVKSLT 375
QY 592 CFLNFPYKIDINVKWKIDGSRQNGVLSWTDQDSDKSTYSMSSTLTTLTKDEYERHNSYT 651
Db 376 CMITDFPEDITVEQWNGQPAEN-YKNTQPIMDT-DGSYFYVSKLVNOKSNWEAGNTFI 433
QY 652 CEATHK 657
Db 434 CSVLHE 439
RESULT 7
US-09-903-327A-12
; Sequence 12, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TAR
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and IGF-1 mature peptide
US-09-903-327A-12
Query Match 29.1%; Score 1023.5; DB 9; Length 510;
Best Local Similarity 37.4%; Pred. No. 2.4e-37;
Matches 249; Conservative 51; Mismatches 111; Indels 255; Gaps 14;
QY 1 EVQLQQSGDPLVKPGASVKISKASGYSTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db 20 EVQLQQSGPELVKPGASVKISKASGYFTDYNMHWKQSHGKSLWIGIYIPYKGGTG 79
QY 61 NQKFKDKATLTVDKSSRAYMELRLTSDSAVYYCARSTMITNTVMYWGOGTSTVYSS 120
Db 80 NQKFKSKATLTDDSSNTAYMELRLTSDSAVYYCARG-----IAYWGGTGLTVTVA 132
QY 121 AKTTPSPVPLAPGSAQAQNSMVTLCCLVKGYPPEPVTVTNWNSGSLSSGVHTFPVAVLQSD 180
Db 133 AKTTPSPVPLAPGSAQAQNSMVTLCCLVKGYPPEPVTVTNWNSGSLSSGVHTFPVAVLQSD 192
QY 181 LYTSSSVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDGGPSEKSEINEKDLRKK 240
Db 193 LYTSSSVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDGG----- 235

Db 193 LYTSSVTPSPSTWSETVTCNVAHPASSTKVKKIVPRDCG----- 235

QY 241 SELQGTALGNLKOIYYNSKAITSSSEKSDAQFLNTLLEKGFTHPWYNDLLVDLGSTA 300

Db 236 CKPCICTVPEVSSVFPPK-----PKDVLITL----- 264

QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360

Db 265 -----TP-KVTCVVVDIS-----KDDPEVQSFWEVD--- 289

QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVORGLIYF 411

Db 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHQDWLNGKEKCRVNSAAPAPIEK----- 343

QY 412 HSSEGSTVSYDLFDAQGYPDTLRLRIYRDNNTTSSLSLSISLYLYTTSIVMTQPTPTSLLV 471

Db 344 -----TISKT----- 348

QY 472 SAGDRVTITCKASQSVNSDVAMVYQKPGOSPCKLLISYTSRYAGVDPDRFSGSGYGTDFTL 531

Db 349 ----- 348

QY 532 TISSVQAEADAAVYFCQDYNPPTFGGTTKLEIKRADAAPTYSIFPPSSEQLTSGGASVY 591

Db 349 -----KGRPKAPQVYTIPTPPKEQMAKDKVSLT 375

QY 592 CFLNFPKIDINVKKIDGSEKQNGVLNSWTQDSDKSTYSMSSTLTLTQDEYERHNSYT 651

Db 376 CMITDFEPEDITVEQWNGQPAEN-YKNTQPIMDT-DGSYFYVYKLNQVKSWEAGNTFI 433

QY 652 CEATHK 657

Db 434 CSVLHE 439

RESULT 8

US-09-903-327A-6

; Sequence 6, Application US/09903327A

; Patent No. US20020164333A1

; GENERAL INFORMATION:

; APPLICANT: Nemerow, Glen R.

; APPLICANT: Li, Erguang

; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET

; TITLE OF INVENTION: GENE

; TITLE OF INVENTION: DELIVERY

; FILE REFERENCE: 22908-1228

; CURRENT APPLICATION NUMBER: US/09/903,327A

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 09/613,017

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 438

; TYPE: PRT

; ORGANISM: Mouse

; FEATURE:

; NAME/KEY: PEPTIDE

; LOCATION: (0)...(0)

; OTHER INFORMATION: Portion of DAV-1 heavy chain used for fusion protein

US-09-903-327A-6

Query Match 29.0%; Score 1022.5; DB 9; Length 438;

Best Local Similarity 37.4%; Pred. No. 2.3e-37;

Matches 249; Conservative 50; Mismatches 111; Indels 255; Gaps 14;

QY 1 EVLOQSGPDLVKPGASVKISKASGYSTGYTMHWKSPGKGLIEWIRNPNGVTLY 60

Db 20 EVLOQSGPELVKPGASVKISKASGYTFTDYNMHWKSHGKSLIEWIGYIYKGGTCY 79

QY 61 NQKFKDKATLTVDKSTTAYMELRSLTSEDSAVYYCARSTMTITNYYWVGOGTSTVSS 120

Db 80 NQKFKSKATLTDTSSSNTAYMELRSLTSDASAVYYCARG-----IAYWGGTTLTVSA 132

QY 121 AKTTTPSVYPLAPGSAQTNSMVTLCGLVKGYFPEPVTVTWNSGSLSSGVHTFPVAVLQSD 180

Db 133 AKTTTPSVYPLAPGSAQTNSMVTLCGLVKGYFPEPVTVTWNSGSLSSGVHTFPVAVLQSD 192

QY 181 LYTSSVTPSPSTWSETVTCNVAHPASSTKVKKIVPRDGGPSEKSEINEKDLRKK 240

Db 193 LYTSSVTPSPSTWSETVTCNVAHPASSTKVKKIVPRDCG----- 235

QY 241 SELQGTALGNLKOIYYNSKAITSSSEKSDAQFLNTLLEKGFTHPWYNDLLVDLGSTA 300

Db 236 CKPCICTVPEVSSVFPPK-----PKDVLITL----- 264

QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360

Db 265 -----TP-KVTCVVVDIS-----KDDPEVQSFWEVD--- 289

QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVORGLIYF 411

Db 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHQDWLNGKEKCRVNSAAPAPIEK----- 343

QY 412 HSSEGSTVSYDLFDAQGYPDTLRLRIYRDNNTTSSLSLSISLYLYTTSIVMTQPTPTSLLV 471

Db 344 -----TISKT----- 348

QY 472 SAGDRVTITCKASQSVNSDVAMVYQKPGOSPCKLLISYTSRYAGVDPDRFSGSGYGTDFTL 531

Db 349 ----- 348

QY 532 TISSVQAEADAAVYFCQDYNPPTFGGTTKLEIKRADAAPTYSIFPPSSEQLTSGGASVY 591

Db 349 -----KGRPKAPQVYTIPTPPKEQMAKDKVSLT 375

QY 592 CFLNFPKIDINVKKIDGSEKQNGVLNSWTQDSDKSTYSMSSTLTLTQDEYERHNSYT 651

Db 376 CMITDFEPEDITVEQWNGQPAEN-YKNTQPIMDT-DGSYFYVYKLNQVKSWEAGNTFI 433

QY 652 CEATH 656

Db 434 CSVLH 438

RESULT 9

US-09-870-759-8

; Sequence 8, Application US/09870759

; Patent No. US20020177551A1

; GENERAL INFORMATION:

; APPLICANT: TERMAN, David S

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE

; FILE REFERENCE: 870759

; CURRENT APPLICATION NUMBER: US/09/870,759

; CURRENT FILING DATE: 2002-01-14

; PRIOR APPLICATION NUMBER: US 60/208,128

; PRIOR FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 166

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 257

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-870-759-8

Query Match 26.5%; Score 935; DB 9; Length 257;

Best Local Similarity 75.5%; Pred. No. 7.7e-34;

Matches 176; Conservative 21; Mismatches 36; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSSSEKSDAQFLNTLLEKGFTHG 285

Db 25 SEKSEINEKDLRKKSELOGTALGNLKOIYYNEKAKTENKESHQOFLQHTILFKGFFTN 84

QY 286 HPWYNDLLVDLGSTAATSEYSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345

Db 85 HSWYNDLLVDFSDKDIVDKYKGVKVDLYGAYGYQACGTPNKTCACMGVTVLHDNNRLT 144
Qy 346 EKKKVPINLWIDGKOTVPIDKVTKSKKEVTVOELDLQARHVLHGKFGLYNSDSFGGKVQ 405
Db 145 EKKKVPINLWIDGKONTVPLETVKTNKKNVTVOELDLQARRYLQKYNLYNSDVFDDKVQ 204
Qy 406 RGLIVFHSSEGSVSYDLFADGQYPTDLLRIYRDNNTTISSTLSLSLYLXTT 458
Db 205 RGLIVFHTSTEPSVNYDLFGAQQNSNTLLRIYRDNKNTINSENMMHIDIVLYTS 257

RESULT 10
US-09-903-327A-4
; Sequence 4, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erquang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22508-1238
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (0)...(0)
; OTHER INFORMATION: DAV-1 light chain, penton base monoclonal antibody
US-09-903-327A-4

Query Match 25.9%; Score 912; DB 9; Length 238;
Best Local Similarity 81.3%; Pred. No. 6.8e-33;
Matches 178; Conservative 9; Mismatches 28; Indels 4; Gaps 1;

Qy 457 TTSIVMTQPTSLVYSGDRVTITCKASQVSDND---VAVYQQRPGQSPKLLISYTSR 512
Db 19 TGDIVLTQSPASLVSLGQRATISCKASQSDVDGDSYMNWYQQRPGPKLLIYAASNL 78
Qy 513 YAGVDRSGSGYGTDFLTLSVQAEADAAVFCQDYNPPTFGGKTKLEIKRADAAPT 572
Db 79 ESGIPARFSGSGGTDFTLNIPHEVEEDAATYICQQTNEPDWTFGGGKTKLEIKRADAAPT 138
Qy 573 VSIFPPSSQLTSGGASVVCFLNNFYPKDINVKWKIDGSEKNGVLSNWTDDSDKSDSTYS 632
Db 139 VSIFPPSSQLTSGGASVVCFLNNFYPKDINVKWKIDGSEKNGVLSNWTDDSDKSDSTYS 198
Qy 633 MSTLTTLTKDEYRHNSYTCEATHKSTSTSPIVKSFNRRNE 671
Db 199 MSTLTTLTKDEYRHNSYTCEATHKSTSTSPIVKSFNRRNE 237

RESULT 11
US-09-910-059-131
; Sequence 131, Application us/09910059
; Patent No. US20020142359A1
; GENERAL INFORMATION:
; APPLICANT: Copley, Clive G
; APPLICANT: Edge, Michael Derek
; APPLICANT: Emery, Stephen Charles
; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said Antibody,
; TITLE OF INVENTION: Their Therapeutic use in an Adept System
; FILE REFERENCE: 1991-209
; CURRENT APPLICATION NUMBER: US/09/910,059
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 09/171,945

; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: PCT/GB97/01165
; PRIOR FILING DATE: 1997-04-29
; PRIOR APPLICATION NUMBER: GB 9703103.3
; PRIOR FILING DATE: 1997-02-14
; PRIOR APPLICATION NUMBER: GB9609405.7
; PRIOR FILING DATE: 1996-05-04
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 131
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: full-length human B7.1-murine ASB7 Fd fusion
US-09-910-059-131

Query Match 24.3%; Score 855.5; DB 10; Length 473;
Best Local Similarity 72.9%; Pred. No. 3.1e-30;
Matches 164; Conservative 24; Mismatches 34; Indels 3; Gaps 3;

Qy 1 EVOLQSGPDLVKPCASVKISKASGYSFTGYIMHWKQSPCKGLEWIGRI-NPNNGVTL 59
Db 243 EVKLVEGGGLVQPGGSLRLSCATSGFTTDTYIMNWRQPPGKALEWLGFIGNKANGYTT 302
Qy 60 -YNOKFKDKATLTVDKSTTAYMELRSLTSDSAVYVCARSTMITNYYMDYWGQGTSTVTV 118
Db 303 EYSASVKGRFTISRDKSOSILYLQNTLRAEDSAIYYCTRDRGLRFY-FDWGQGTTLTV 361
Qy 119 SSAKTPPSVYPLAPGSAATNSMVTGLGKLVKGYPEPPTVTWNSGSLSSGVHTFPVLQ 178
Db 362 SSAKTPPSVYPLAPGSAATNSMVTGLGKLVKGYPEPPTVTWNSGSLSSGVHTFPVLQ 421
Qy 179 SDLYTLSSSVTPSPSTWPTSETVTCNVAHPASSTKVDDKIVPRDSG 223
Db 422 SDLYTLSSSVTPSPSTWPTSETVTCNVAHPASSTKVDDKIVPRDCG 466

RESULT 12
US-09-917-410-6
; Sequence 6, Application US/09917410
; Patent No. US20020098183A1
; GENERAL INFORMATION:
; APPLICANT: MARTIN, Ulrich; HASELBECK, Anton; SCHUMACHER, Guenther;
; CO. Man S.
; TITLE OF INVENTION: ANTI-L-SELECTIN ANTIBODIES FOR PREVENTION OF
; MULTIPLE ORGAN FAILURE AFTER POLYTRAUMA AND FOR
; PREVENTION OF ACUTE ORGAN DAMAGE AFTER
; EXTRACORPOREAL BLOOD CIRCULATION
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Computer Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII, Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/917,410
; FILING DATE: 26-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/578,953
; FILING DATE: <Unknown>
; APPLICATION NUMBER: EP 95 114 969.9
; FILING DATE: 19-Sep-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020098183A1man D.

```

;
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: BOER 1059-PFF/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 448 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-917-410-6

Query Match      22.9%; Score 807; DB 10; Length 448;
Best Local Similarity 30.8%; Pred. No. 3.4e-28;
Matches 211; Conservative 67; Mismatches 126; Indels 280; Gaps 18;

Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYIMHWKQSPGKGLWIGRINPNNGVTLY 60
Db 1 QVQLVQSGAEVKKPGSSVKVCKASGYFTSYIMHWVRQAPGQGLEWIGYIYPYNDGTY 60

Qy 61 NQKFKDKATLTVDKSSSTAYMELSLTSEDSAVYYCARSTMITNVV--MDYWGQGTSTVTV 118
Db 61 NEKFKGRVTIISDESTNAYMELSLRSEDYAVYICAREE-YGNIVRVFEDVWGQGLTVTV 119

Qy 119 SSAKTPPSVYPLAPGSAOATNSMTVLGLVKGYPEPVPVTVTNWSSGSLSSGVHTFPAVLQ 178
Db 120 SSASTKGPSVPLAPCSRSTSESTAALGCLVKDYFPEPVTVSNWNSGALTSGVHTFPAVLQ 179

Qy 179 SD-LYTLSSVTVSPSTWSPSETVTCNVAPASSTKVDKKI-----VPRDSGGP 225
Db 180 SSGLYSLSSVTVSPSSLTGTYTCNVDPKPSNTKVDKRVESKYGPCPCPAPEFLGGP 239

Qy 226 SEKSEEINEKDLRKSELQGLTALGNLKOIYYNYSKAITSSSEKSAQDFTLTLFKGFTTG 285
Db 240 S-----VELFPPKP-----KDTLMI----- 254

Qy 286 HPWYNDLLVLGSLTAATSEYEGSSVDLYGAYGYQCAGGTGPNKTACMYGGVTLHDNNRLT 345
Db 255 -----SRTPEVT-CVVVDVS-----Q 269

Qy 346 EEKKYPINLWIDGKQTTVPIDKVKTSKEV-----TVQELDLQARHYLHGKFGLYNSD 398
Db 270 EDPEYQFNWYVDG-----VEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGK----- 318

Qy 399 SPGKGVQRLIVFHSSEGSTVSYDLFDAQGOVPTLLRIYRDNTTISSTLSISLVLYTT 458
Db 319 -----EY----- 320

Qy 459 STVMTQTPTSLVLSAGDRVTITCKASQSVNDVAWYQKPGQSPKLLISYTSRYSAGVDP 518
Db 321 -----KCKVSNK-----GLP- 330

Qy 519 RPSGSGYGTDFLTITSSVQAEDAAYVFCQDYNPPTFGGGTKLEIKRADAAPTVSIIPP 578
Db 331 -----SSIEK-----TISKAKQPREPVYITLPP 354

Qy 579 SSEQLTSGGASVVCFLNFPKIDINVKWKIDGSRQN-----GVLNSWTDQDSKDSYSM 633
Db 355 SGEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDLSD-----DCSFFL 407

Qy 634 STLTTLTKDEYERHNSYTCETHK 657
Db 408 YSRLTVDKSRMQEGNVFSCSVNHE 431

RESULT 13
US-09-249-011A-24
; Sequence 24, Application US/09249011A
; Patent No. US20020176855A1
; GENERAL INFORMATION:
; APPLICANT: CO, MAN SUNG
; APPLICANT: VASQUEZ, MAXIMILIANO
```

```

; APPLICANT: CARRENO, BEATRIZ
; APPLICANT: CELNIKER, ABBIE CHERYL
; APPLICANT: COLLINS, MARY
; APPLICANT: GOLDMAN, SAMUEL
; APPLICANT: GRAY, GARY S.
; APPLICANT: KNIGHT, ANDREA
; APPLICANT: O'HARA, DENISE
; APPLICANT: RUP, BONITA
; APPLICANT: VELDMAN, GEERTRUIDA M.
; TITLE OF INVENTION: HUMANIZED IMMUNOGLOBULIN REACTIVE WITH B7-2 AND METHODS
; TITLE OF INVENTION: OF TREATMENT THEREWITH
; FILE REFERENCE: 08702.0081-00000
; CURRENT APPLICATION NUMBER: US/09/249.011A
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 24
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-249-011A-24

Query Match      22.5%; Score 792; DB 9; Length 461;
Best Local Similarity 30.1%; Pred. No. 1.5e-27;
Matches 203; Conservative 73; Mismatches 132; Indels 266; Gaps 15;

Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYIMHWKQSPGKGLWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPGSSVKVCKASGYFTDYAIOVRQAPGQGLEWIGVINIYDNTNY 79

Qy 61 NQKFKDKATLTVDKSSSTAYMELSLTSEDSAVYYCARSTMITNVYMDYWGQGTSTVTVSS 120
Db 80 NQKFKGKATMTVDKSTSTAYMELSLRSEDYAVYYCARAA---WYMDYWGQGLTVTVSS 135

Qy 121 AKTPPSVYPLAPGSAOATNSMTVLGCLVKGYFPPEPVTVTNWSSGSLSSGVHTFPAVLQSD 180
Db 136 ASTKGPSVPLAPCSRSTSESTAALGCLVKDYFPEPVTVSNWNSGALTSGVHTFPAVLQSD 195

Qy 181 -LYTLSSVTVSPSTWSPSETVTCNVAPASSTKVDKKLVPR-----DSGGPSEK 228
Db 196 GLYSLSSVTVSPSNFQTQTYTCNVDPKPSNTKVDKTVVERKCCVCEPCPCPAPPAAPS--- 253

Qy 229 SEETNEKDLRKSELQGLTALGNLKOIYYNYSKAITSSSEKSAQDFTLTLFKGFTTGHWP 288
Db 254 -----VELFPPKP-----KDTLMI----- 267

Qy 289 YNDLLVLGSLTAATSEYEGSSVDLYGAYGYQCAGGTGPNKTACMYGGVTLHDNNRLTBEK 348
Db 268 -----SRTPEVT-CVVVDVS-----HEDP 285

Qy 349 KVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDSFGKGVQRGL 408
Db 286 EVQFNWYVDG-----VEVHNAKTKPRE----- 307

Qy 409 IVFHSSEGSTVSYDLFDAQGOVPTLLRIYRDNTTISSTLSISLVLYTTISVMTQTPTS 468
Db 308 -----EQFNSTFRVSVLTVVHOD----- 326

Qy 469 LLVSAGDRVTITCKASQSVNDVAWYQKPGQSPKLLISYTSRYSAGVYDPDRFSGSGYGT 528
Db 327 -----WLNKGEYK-----KVSNGKLGPAP 345

Qy 529 FTLTITSSVQAEDAAYVFCQDYNPPTFGGGTKLEIKRADAAPTVSIFFPSEQUITSGGA 588
Db 346 LEKTSITKGO-----PREPQVYITLPPREEMTKNQV 377

Qy 589 SVVCFLLNFPKIDINVKWKIDGSRQN-----GVLNSWTDQDSKDSYSMSSTLTLTKE 643
Db 378 SLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDLSD-----DCSFFLYSKLTVDKSR 430

Qy 644 YERHNSYTCETHK 657
Db 431 WQGNVFCSVNHE 444
```

RESULT 14
US-09-828-995B-20
; Sequence 20, Application US/09828995B
; Patent No. US20020165135A1
; GENERAL INFORMATION:
; APPLICANT: Heska Corporation
; APPLICANT: McCall, Catherine A.
; APPLICANT: Tang, Liang A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATED TO CANINE IGG AND CANINE IL-13
; FILE REFERENCE: AL-7
; CURRENT APPLICATION NUMBER: US/09/828,995B
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,874
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/195,659
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-828-995B-20

Query Match 22.4%; Score 789; DB 9; Length 473;
Best Local Similarity 30.4%; Pred. No. 2e-27;
Matches 204; Conservative 73; Mismatches 145; Indels 250; Gaps 15;

Qy 1 EVQLQSGPDLVKPGASVKISKASGYFTGYMHVWQSPGKGLEWIGRINPNNGVTLY 60
Db 20 EVRLVESGGTLVKPGSLKSCVAGSFTFRYSMDWVRQAPGKSLQWVAGIDGTGTSY 79

Qy 61 NQKFKDKATLVKDSSTAYWELSLTSDSAVYVCARSTMITNYVMYWGQGTSTVVS 120
Db 80 SOTVGRFTISRDNAKNTLYLQINSRAEDSAVYCAKS--WSRNGDLWVGQGLTVTVSS 138

Qy 121 AKTTPSYVPLAPGSAATNSMTLGLCLVKGYFPEPVTVTVNSGSLSSGVHTFPAVLQSD 180
Db 139 ASTTAPSVPLAPSGSTSGSTVALACLVSGLYFPEPVTVSNWSGSLTSGVHTFPAVLQSD 198

Qy 181 -LYTLSSSVTPSPSTWTPSETVTCNVVAHPASSTKVDKKI-----VPRDSGGSPSKSEI 232
Db 199 GLYSLSSVTPSPSSRWSPSETETCNVAHPASTKVDKVPKRENGRVRPRPDCPKPAPEM 258

Qy 233 NEKDLRKKSELQGTALGNLKOIYYNSKAITSSSEKSAQDLTNTLLFKGFTGHPWYNDL 292
Db 259 -----LGG-PSVFIIPPKP-----KDTLLI----- 277

Qy 293 LVDLGSTAATSEYEGSSVDLYGAYGYOCAGTGNKTAACMYGGVTLHDNNRLTEKKVPI 352
Db 278 -----ARTPEVTCVVVDL-----PEDPEVQI 299

Qy 353 NMDWIDGKQTTPIDKVKTSKEV-----TVQELDLQARHYLHGKFGLYNSDSFGKVK 405
Db 300 SFWFDGKQ-----MQTAKTPREEQNGTYRVVSVLPVIGHQDWLKGK----- 341

Qy 406 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTLSLSLYLTTSIVMTQT 465
Db 342 ----- 341

Qy 466 PTLSSLVASGRDVTITCKASQSVNDVAVYQKPGQPKLLISYTSRRYAGVPDRFSGSY 525
Db 342 -----QFTCKVN-----NKALPSP---IERTISKARG----- 365

Qy 526 GTDFTLTISVQAEADAVYFCQDYNPSPTGGGKLEIKRADAAPTYSIIFPPSSEQLTS 585
Db 366 -----QAHPSVY-----VLPPSREELSK 384

Qy 586 GGASVVCFLNFPKDIYKWKIDSERQNGVLSNWTDDSKDSTYSMSSTLTLTCKDEYE 645
Db 385 NTVSLTCLIKOFFPPDIDVWQSNQEQEPESKYRTTPQLDDEGSGFYFLYSKLSVDKSRWQ 444

Qy 646 RHNSYTCETHK 657
Db 445 RGDTFICAVMHE 456

RESULT 15
US-09-825-012-52
; Sequence 52, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-52

Query Match 22.4%; Score 789; DB 10; Length 729;
Best Local Similarity 31.5%; Pred. No. 2.9e-27;
Matches 209; Conservative 65; Mismatches 151; Indels 238; Gaps 17;

Qy 1 EVQLQSGPDLVKPGASVKISKASGYFTGYMHVWQSPGKGLEWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPKASVKVCKASGYFTSAIWIWVRQAPGKGLWYGEILPGSNRSY 79

Qy 61 NQKFKDKATLVKDSSTAYWELSLTSDSAVYVCARSTMITNYVMYWGQGTSTVVS 120
Db 80 NEKFKGRVTVTRDSTNTAYWELSLRSEDATVYVCARSYDFANFA--YWGQGLTVTVSS 137

Qy 121 AKTTPSYVPLAPGSAATNSMTLGLCLVKGYFPEPVTVTVNSGSLSSGVHTFPAVLQSD 180
Db 138 ASTKGPSVFEPLAPSGKSTSGGTAALCLVKDYFPEPVTVSNWSGALTSGVHTFPAVLQSS 197

Qy 181 -LYTLSSSVTPSPSTWTPSETVTCNVVAHPASSTKVDKKIIPRDSGGSPSKSEINEKDLRK 239
Db 198 GLYSLSSVTPSPSSSLGTQTYICNVNHNKPSNTKVDKKYEPKSC----- 240

Qy 240 KSELQGTALGNLKOIYYNSKAITSSSEKSAQDLTNTLLFKGFTGHPWYNDLLVLDLGS 299
Db 241 -----DKTHYCPCPAPELGGPSVF--LFPKP--KDTLM----- 272

Qy 300 AATSEYEGSSVDLYGAYGYOCAGTGNKTAACMYGGVTLHDNNRLTEKKVPIINLWIDGK 359
Db 273 -----ISRTPEVT-CVVVDVS-----HEDPEVKFNWVDG- 301

Qy 360 QTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDSFGKVGQGLIVFHSSEGSTV 419
Db 302 ---VEVHNKATKPREEQ-----YNSTY---RVVSVLTVLHQDWLNGK 337

Qy 420 SYDLFDAGQYPTDLLRIYRDNNTTISSTLSLSLYLTTSIVMTQTPTSLLSVASGRDVTI 479
Db 338 EY----- 339

Qy 480 TCKASQSVNDVAVYQKPGQPKLLISYTSRRYAGVPDRFSGSYGDTFTLTISVQAE 539
Db 340 KCKVS-----NKALPAP---IEKTSKAGQD----- 363

Qy 540 DAAYVFCQDYNPSPTGGGKLEIKRADAAPTYSIIFPPSSEQLTSGASVVCFLNFPY 599
Db 364 -----REFQVYTLPPSRDELTKNQVSLTLYVKGEYP 394

Query Match	22.4%	Score 789;	DB 10;	Length 731;
Best Local Similarity	31.5%	Pred. No. 2.9e-27;		
Matches 209;	Conservative 65;	Mismatches 151;	Indels 238;	Gaps 17;
Qy	1	EVOLQSGPDLVPCASVKISKCKASGYSTGYHHWVKQSPGKLEWIGRINPNCVTLY	60	
Db	20	QVQLVQSGAEYKPKASVKYCKASGYTFSYAIIEWVQAPKGLWYGEILPGSNRRY	79	
Qy	61	NOKRDKATLIVDKISSTYAYWELKSLTSEDSAVYVCARSTITNYVMYWGQGSTVTVSS	120	
Db	80	NEKFKGRVTYTRDTSTNTAYWELSLRSEDYAVYCAKSYDFANFA--YWGQGLTVTVSS	137	
Qy	121	AKTTPPSVYPLAPGSAAGTNSMTVIGCLVKGYFPEPVTVTWNSGSLSSGVHTFPVAVLQSD	180	
Db	138	ASTKGPSVFPLAPSSKPTSGCTAAALGCLVKDYFPEPVTVSNNSGALTSGVHTFPAVLQSS	197	
Qy	181	-LYTLSSSVTVPSSTWBPSETVTCNVAHPASSTKVDKDKLVPRDSGGPSEKSEINEKDLRK	239	
Db	198	GLYSLSVVTVPSSSLGTQTICNVNHRKPSNTKVDKKVEPKSC-----	240	
Qy	240	KSELOGTALGNLKQIYYNNSKAITSSKSAQDLTNTLLFKGFTGHPWYNDLLVDLGST	299	
Db	241	-----DKTHTCPCPAPDELGGPSVF--LFPKPK--KDTLM-----	272	
Qy	300	AATSEYEGSSVDLYCAYGYCAGGTPNKTKACMYGGVTLHDNNRLTEKKKYPINLWIDGK	359	
Db	273	-----ISRTPEVT--CVVVDVS-----HEDPEVKFNWYVDG--	301	
Qy	360	QTTVPIDKVKTSKKEVTVQEELDLQARHYLHGKFLGYNDSFGKVGQRGLIVFHSSSESTV	419	
Db	302	---VEVHNAKTKPREQ-----YNSTY---RVVSVLTVLHQDWLNGK	337	
Qy	420	SYDLFDAQGYQPDTLRLIRYRDNNTISSTLSISLYTTSIVMTQPTPSLVLVSAGDRVTI	479	
Db	338	EY-----	339	

```

Qy 480 TCRASQSVSNDVAWYQQKPGQSPKLLISYTSSRYAGVDPDRFGSGYGDTFTLTITSSVQAE 539
    || |      | : | | | : | |
Db 340 KCKVS-----NKALPAP---IEKTISKAKGP-----363
Qy 540 DAAYVFCQDYNPPPTGGGTTKLEIKRADAAPTVISFPSPSEQLTSGGASVVCFLNMFYP 599
    | | | | | | | | | | | | | | | | | | | | | |
Db 364 -----REPOVYTLPPSRDELTKNQVSLTCLVKGFYP 394
Qy 600 KDINVKWKIDGSERQN-----GYLNSWTDDSKDSYSMSSTLTCLKDEVERHNSVTCEA 654
    || | |:| | | | | | | | | | | | | | | | | |
Db 395 SDIAVESMGQPENNYKTPPVLD-----DGSEFLYSKLTVDKSRMQGNVFCSV 447
Qy 655 THK 657
    | :
Db 448 MHE 450
RESULT 18
US-09-825-012-61
; Sequence 61, Application US/09825012
; Patent No. US2002012798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFc1 heavy chain - DNase I fusion
US-09-825-012-61

```

Query Match	22.4%	Score 789;	DB 10;	Length 739;
Best Local Similarity	31.5%;	Pred. No. 3e-27;		
Matches 209;	Conservative 65;	Mismatches 151;	Indels 238;	Gaps 17;
Qy	1	EVQLQQSGPDLVKGASVKISCKASGYSFTCYGMHWKVKQSPKGLEWIGTRINPNNGVTLY	60	
Db	20	QVQLVQSGAEVKKFGASVKVSKASGYSFTSAIWEHWKQAPKGLEWVEIGLPGSNRY	79	
Qy	61	NQKPKDQATLTVDKSSSTAYMELRSLTSEDSAVYVCARSTMITNVMYDQGGTSVTYSS	120	
Db	80	NEKEFGKRVTVTRDTSTNTAYMELSLRSEDFAVYVCARSYDFAWFA--YWGQGLTVTVSS	137	
Qy	121	AKTTPPSVYPLAPGSAQGTNSWVLGCLVKCYGPEPPTVTVWNSGSSLSGCVHTFPAVLQSD	180	
Db	138	ASTGKSVFPLAPGSKSTSGGTAALGCLVKDYFPEPPTVSNWNSGALTSGVHTFPAVLQSS	197	
Qy	181	-LYTLSSSVTVPSSTWPSSETVCNVAHPASSTKVDKKIIVPRSDSGSPSEKSEINEKDLRK	239	
Db	198	GLYSLSVTVTVPSSSLGTQTICVNNHKPSNTKVDKKEPKSC-----	240	
Qy	240	KSELOGTALGNLKOIYYNKAITSSEKSAQOFLTNTLLFKGFTGHPWYNDDLVDLGST	299	
Db	241	-----DKTHTCPPCPAPELIGGPSVF--LFPKP--KDTLM-----	272	
Qy	300	AATSEYEGSSVDLYGAYGYCAGGTGPKTKACMGVGTVLDHNNRLTEKKKVPINLWIDGK	359	
Db	273	-----ISRTPEVT-CVVVDVS-----HEDEPKFKNWYDGC-	301	
Qy	360	QTTVPIDKVKTSKKEVTVQVLELDQARHYLHGKFLGYNSSDFSGKQVQRGLIVFHSSEGTV	419	
Db	302	---VEVHNAKTKPREQ-----YNSTY---RVVSVTLVHODLWNGK	337	

Qy	420	SYDLFDAQGVDPDLLRIYRDNMTTSSLSLSISLYLTYTTSIVMTQTPTSLLVSAGDRVTI	479
Db	338	EY-----	339
Qy	480	TKCASQSVSNDVANYQKPGOSP KLLISYTSR YAGVPDRFSGSGYGTDTLTITISSVQAE	539
Db	340	KCKVYS-----NKALPAP---TEKTI SKAGQP-----	363
Qy	540	DAAVYFCQDYNSPPTFGGKTLEIKRADAAPTVSIFFPSSEQLTSGCASVVCFLNNFY	599
Db	364	-----REPOVYTLPPSDELTKNQVSLTCLVKGFYP	394
Qy	600	KDINVKWKIDGSEQRN-----GVLSNMTDQSKDSTYSMSSTLTLTKDEYERHNSYTC	654
Db	395	SDIAVEWESNGQPENNYKTTTPVLDSD-----DGSFFLYSKLTVDKSRWQGNVFC	447
Qy	655	THK 657	
Db	448	MHE.450	
RESULT 19			
US-09-825-012-58			
; Sequence 58, Application US/09825012			
; Patent No. US20020122798A1			
; GENERAL INFORMATION:			
; APPLICANT: Young, Robert			
; TITLE OF INVENTION: Compounds for Targeting			
; FILE REFERENCE: 43191-256808			
; CURRENT APPLICATION NUMBER: US/09/825.012			
; CURRENT FILING DATE: 2001-04-03			
; PRIOR APPLICATION NUMBER: US 60/237,159			
; PRIOR FILING DATE: 2000-10-02			
; PRIOR APPLICATION NUMBER: GB 0008049.9			
; PRIOR FILING DATE: 2000-04-03			
; NUMBER OF SEQ ID NOS: 102			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 58			
; LENGTH: 740			
; TYPE: PRT			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion			
US-09-825-012-58			
Query Match 22.4%; Score 789; DB 10; Length 740;			
Best Local Similarity 31.5%; Pred. No. 3e-27;			
Matches 209; Conservative 65; Mismatches 151; Indels 238; Gaps 17;			
Qy	1	EVQLQQSGPDLVPCGASVKISKASGYSFTGYVHHWKQSPGKLEWIGRINPNNGVTLY	60
Db	20	QVQLVQSGAEVKVPCGASVKYKASGYSFTSAWIEWRQAPGKLEWVGETLPGSNNSRY	79
Qy	61	NOKEFKATLTVDKSSITAYMELASLTSEDSAVYVCARSTMITNYVMQYSGQSTVTSS	120
Db	80	NEKFKGRVTVTROISTNTAYMELSSLSRSEDFAVYICAKSYDFANFA--YWQGGTLFVTVSS	137
Qy	121	AKTTPPSYIPLAPGSAQNTSMVTLGCLVKGYFPEPVTTVWNSGSLSSGVTHTFPALQSD	180
Db	138	ASTGKPSVFELAPSSKSTSGGTAALGCLVKDYFPEPVTVSNNSGALTSGVHTFPAVLQSS	197
Qy	181	LYTLSSVTVPPSWTWSFETVTCNVAHPASTKVDKLVIPRDSGSPSEKSEINEKDLRK	239
Db	198	GLYSLSSVTVPPSSLSLGTQTICNVNHPKPSNTKVDKVEPKSC-----	240
Qy	240	KSELQGTALGNLKQIIYYVNSKAITSSEKSAQDELNTLTLLFKGFTGHPWYNDLLVLDLGS	299
Db	241	-----DKTHTCCPCPAPELLGGSVF--LFPKP--KDTLM-----	272
Qy	300	AATSEYEGSSVDLYGAYYGOCAGTGNKTKACMYGGVTLHDNNRLTEBKKVPINLWIDGK	359
Db	273	-----ISRTPEVT--CVVDVYS-----HEDPEVKFNWYVDG--	301

	QY	360	QTTPIDKVTSKKEVTVQEELDLQAIRHYLHGKFGLYNSDSFSGKVQRGLIVFHSEGSTV	419
	Dg	302	--VEVNNAKTPREEQ-----YNSTY---RVSVSLTVLRHQDWLNKG	337
	QY	420	SYDLFDAQGOYPDPLLLRIYRDNTTISSTSLSLSLYLTTSIVMTOTPTSLVSAGDRVTI	479
	Dg	338	EY-----	339
	QY	480	TCKASQSVDVAWYOQKPQGSPKKLLIYSTSSRYAGVPDRFSGSGYGDTFLTITISSVOAE	539
	Dg	340	KCKVS-----NKALPAP--TEKITSAKGGQP-	363
	QY	540	DAAVFYCCQDYSNPPTFGGTGKLKRADAAPTVISIPPPSSEQLTSGASGVVCFLNNPY	599
	Dg	364	-----REPOVYTLPSPSRDELTKNQVSLCWLVKGFYP	394
	QY	600	KDINVWKIKDGSERQN----GVLSNWTDOOSKDSTYSMSSTLTTLTKDEXERHNSYTCEA	654
	Dg	395	SDATAVENESGNPENNYKTTTPPVLDSD-----DGSFFLYSKLVDKRWOOGNVFSCSV	447
	QY	655	THK 657	
	Dg	448	MHE 450	
 RESULT 20 US-09-825-012-55 ; Sequence 55, Application US/09825012 ; Patent No. US20020122798A1 ; GENERAL INFORMATION: ; APPLICANT: Young Robert ; TITLE OF INVENTION: Compounds for Targeting ; FILE REFERENCE: 43191-256808 ; CURRENT APPLICATION NUMBER: US/09/825_012 ; PRIOR FILING DATE: 2001-04-03 ; PRIOR APPLICATION NUMBER: US 60/237,159 ; PRIOR FILING DATE: 2000-10-02 ; PRIOR APPLICATION NUMBER: GB 0008049.9 ; PRIOR FILING DATE: 2000-04-03 ; NUMBER OF SEQ ID NOS: 102 ; SOFTWARE: Patentin version 3.1 ; SEQ ID NO 55 ; LENGTH: 741 ; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNaase I fusion US-09-825-012-55				
		Query Match	22.4%; Score 789;	DB 10; Length 741;
		Best Local Similarity	31.5%; Pred. No.	3e-27;
		Matches	209; Conservative	65; Mismatches 151; Indels 238; Gaps
	QY	1	EVQLQQSGPDILVRPGASKVSKCASGYSFCTGYMHMWKQSPGKGLEWIGRINPNNGVTLY	60
	Dg	20	QVLVQSSAEAKPKGASVKYCSKASGYTF SAYIEWRWQARQPKGLEWCETILPGSNRSY	79
	QY	61	NQAFKOKATLTVDKSSITYAYMELRLSYEGEASVVYYCARSTMITYMDYWGWCGTSVTYSS	120
	Dg	80	NEXFKGRVTVTTRDTSTNTAYMELSSLASEDPATVYCARSYDFAWFA--YWGQGLTVTVSS	137
	QY	121	AKTTPPSVYPLADGAOAQTNMWTLCGLVKCYFPPEPTYVTWNSSSLSSGCVHTFPVALQSD	180
	Dg	138	ASTKGPSSFVPLAPSSKSTSGTAALGCLVKDYFPPEPTYVTWNSSGALTSGVHTFPVALQSS	197
	QY	181	-LYTLSSSVTPSPSTWPSETVTCNHAPHASPSTKYDDKIIVPRDMSGGPGSEKSBEINEKDLRK	239
	Dg	198	GLYSLSSVTVTPSSSLTGTYICNVHNKPSNTKYDKKVEPKSC-----	240
	QY	240	KSELQGTALGNLKIOIYYNNSKAITSSEKSAQOFUTNTLLFGFTGHHPWINDLLVDLGST	299
	Dg	241	-----DKTHFCPCCPAPELLGGSVF--LEPPKP--KDTLM-----	272

```

Qy 300 AATSEYGSSVDLYGAYYGYQCAGGTPNKATACMYGGVTLHDNNRLTEKKKVPINLWIDGK 359
Db 273 -----ISRTPEVT-CVVVDVS-----HEDPEVKFNYYVDG- 301
Qy 360 QTTYPIDKVKTSKKEVTVQELDLQARHVLHGKFGLYNSDSFGGKQVORGLIVPHSSEGSTV 419
Db 302 ---VEVHNAKTPREEQ-----YNSTY---RVVSVLTVLHQDWLNGK 337
Qy 420 SYDLFDLAQOGQPDTLRLIYRONTTISSTSLISLYLTTSIVMTQTPTSLLYSAGDRVTI 479
Db 338 EY----- 339
Qy 480 TCKASQSVSNDVAVYQKPGOSPGLLIYSYTSRYAGVDPDRFSGSGYGTDTFTLTISVQAE 539
Db 340 KCKVS-----NKALPAP---IEKTSKAKGP----- 363
Qy 540 DAAVYFCOODYNSPPTFGGGTKLEIKRADAPTVSIFPPSSPOLTSGGASVVCFLNNEYP 599
Db 364 -----REPQVYTLPPSRDELTKNQVSLTCLVKGYGYP 394
Qy 600 KDINVKWKIDGERON-----GVLSNWTDDQSDKSTYSMSSTLTTLTKDEYERHNSYTGCEA 654
Db 395 SDIAVEMESNGCPENNYKTTPEVLDS-----DGSFFELYSKLTVDKSRWQQGNVFCSCV 447
Qy 655 THK 657
Db 448 MHE 450

RESULT 21
US-09-859-053-28
; Sequence 28, Application US/09859053
; Patent No. US20020102658A1
; GENERAL INFORMATION:
; APPLICANT: Tsuji, Takashi
; APPLICANT: Tezuka, Katsunari
; APPLICANT: Hori, No. US20020102658A1uaki
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A
; TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE AILIM AND
; TITLE OF INVENTION: PHARMACEUTICAL USE THEREOF
; FILE REFERENCE: 06501-079001
; CURRENT APPLICATION NUMBER: US/09/859,053
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: JP 2001-99508
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: JP 2000-147116
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-859-053-28

Query Match 22.3%; Score 785.5; DB 10; Length 470;
Best Local Similarity 29.6%; Pred. No. 2.9e-27;
Matches 201; Conservative 75; Mismatches 136; Indels 267; Gaps 15;

Qy 1 EVQLQQSGPDLVKPGASVKISCKASGYSTFTGYMHMWKQSPCKGLEWIGTRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPGASVKVCKASGYTFTGYMHMWKQAPQGQGLEWMGWINPHSGGTNY 79
Qy 61 NQKDKDKATLTVDKSSSTTAYMELRSLTSEDSAVYYCAR-----STMITNIVMYDYGQGTFS 115
Db 80 AQKPGQRTVMTDRDTSISTAYMELSLRLSDDTAVYYCARTYYTSSGYYHDAFDINGQGTM 139
Qy 116 VIVSSAKTTPSPVYPLAPGSAAOITNSMTVLCLVKGYGPEPVTVTWNSGSLSSGVHTTTPA 175
Db 140 VTVSSAKTSPGVFPFLAPCPSRSTSESTAALQCLVKDYFPEPVTVSWNSGALTSGVHTTTPA 199
Qy 176 VLQSD-LYTLSSSVTVPSSTWPESTVTCNVAHPASSTKVDKKIVPR-----DSG 223

```

Db 200 VLOSGSLVSSVTVFSSNFGTQTCNVHDHKSNTKVDKTVKCKCECCPCPPAPVA 259
Qy 224 GPSEKSEINEKDLKKSELGALGNLKIYYNSKAITSSSEKSAQDLNTLLKGF 283
Db 260 GPS-----VFLPPRP-----KDTLMI----- 276
Qy 284 TGPWYNLLVLDGSTAATSEYSSVDLYGAYYQYQAGTGNPKTACMYGGVTLHDNR 343
Db 277 -----SRTPEVT-CVVVDVS----- 290
Qy 344 LTEKKVPINLWIDGKOTVPIDIKVYTSKKEVTVVOELDLQARHYLHGKFGLYNSDSFGK 403
Db 291 -HEDPEVQFNMYDQ-----VEVHNAKTPRE----- 316
Qy 404 VQGLIVFHSSEGSTVSYDLFDAGQYDPTLLRLIYRDNNTTISSTLSLSLYLTTSVMT 463
Db 317 -----EQNSTFRVSVLTVVHQD----- 335
Qy 464 QTPTSLVLSAGDRVTITCKASQSVSNDVAVTQKPGQSPKLLISYTSRYAGVDPDRFSGS 523
Db 336 -----WLNKREYKC-----KYSNK 349
Qy 524 GYGDFTLTSSVOAEDAAYVFCODYNSPPTFGGKLEIKRADAPTIVSIFPPSSEQL 583
Db 350 GLPAPIEKTISKTKGQ-----PREPQVYTLPPSREEM 381
Qy 584 TSGCASVVCFLNNFYKPDINVKWKIDGERON-----GVLNSWTDQDSKOSTYSMSSTLT 638
Db 382 TKNOVSLTKVKGYPDSIDIAVEWSNGQPNENNYKTTTPMLDS-----DGSFFLYSKLT 434
Qy 639 LTKDEYERHNSYTCATHK 657
Db 435 VDKSRWQGNVFCVSMHE 453

RESULT 22
US-09-995-693-2
; Sequence 2, Application US/09995693
; Patent No. US20020136721A1
; GENERAL INFORMATION:
; APPLICANT: Schwall, Ralph H.
; Tabor, Kelly H.
; TITLE OF INVENTION: Hepatocyte Growth Factor Receptor
; Antagonists and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/995,693
; FILING DATE: 29-No. US20020136721A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/952,235
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0938P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-995-693-2

Query Match 22.1%; Score 780; DB 10; Length 230;
Best Local Similarity 67.7%; Pred. No. 2.7e-27;
Matches 151; Conservative 28; Mismatches 38; Indels 6; Gaps 3;

Qy 1 EVQLQSGPDLVLRPGASVKISKASGYSTGYMHWYKQSPGKLEWIGRINPNNGVTLY 60
Db 1 ZVQLQSQGPDLVLRPGASVKMSKASGYSTFTYMLHWYKQSPGKLEWIGRINPNNGVTLY 60
Qy 61 NQPKDKATLTVDKSSSTAYMELRSLSSEDSAVYICARSTMITNYV--MDYWGOGTSVTV 118
Db 61 NPVKDKATLTVDKSSSTAYMELRSLSLSLTSADSAYICA---TYGSYVSPLDYWGOGTSVTV 117
Qy 119 SSARKTTPSYVPLAPGSAAGTNSMVTGLCLVKGYPFPEPTVTNNSGSLSSGVTHTFPAVLQ 178
Db 118 SSASTKGPSVPLAPSSKSTSGGTAALGCLVKGYPFPEPTVTNNSGSLSSGVTHTFPAVLQ 177
Qy 179 SD-LYTLSSSVTPSPSTWPSSETVTCNVAPASSTKVDKKIVPR 220
Db 178 SSGLYSLSSVTVPSSSLGTQTYICNVNHHKPSNTKVDKKVEPK 220

RESULT 23
US-09-747-669-3
; Sequence 3, Application US/09747669
; Patent No. US20020122807A1
; GENERAL INFORMATION:
; APPLICANT: Dan, Michael D.
; APPLICANT: Saleh, Mansoor
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS, DESIGNATED
; TITLE OF INVENTION: 4B5 THAT SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES
; TITLE OF INVENTION: ENCODING THE FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS
; FILE REFERENCE: 316082001001
; CURRENT APPLICATION NUMBER: US/09/747,669
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 09/111,286
; PRIOR FILING DATE: 1998-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-09-747-669-3

Query Match 22.1%; Score 779.5; DB 10; Length 476;
Best Local Similarity 31.0%; Pred. No. 5.2e-27;
Matches 208; Conservative 69; Mismatches 150; Indels 243; Gaps 17;

Qy 1 EVQLQSGPDLVLRPGASVKISKASGYSTGYMHWYKQSPGKLEWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPGASVKISKASGYSTFTSFDLNNYRQAPGQGLWGMWNPNSGKTGY 79
Qy 61 NQPKDKATLTVDKSSSTAYMELRSLSSEDSAVYICAR-----STMITNYVDYWGOG 113
Db 80 AQFQGRVTRWTRNTSIRIAYMELSGLRSEDATVAVFCARNADNVEMAAIYHYGHDVWGOG 139
Qy 114 TSVTVSSAKTTPSPVPLAPGSAAGTNSMVTGLCLVKGYPFPEPTVTNNSGSLSSGVTHTF 173
Db 140 TTVTVSSASTKGPSVPLAPSSKSTSGGTAALGCLVKGYPFPEPTVTNNSGSLSSGVTHTF 199
Qy 174 PAVLQSD-LYTLSSSVTPSPSTWPSSETVTCNVAPASSTKVDKKIVPRDGGPSEKSEEI 232
Db 200 PAVLQSGSLVSSVTVPSSSLGTQTYICNVNHHKPSNTKVDKKVEPKSC----- 249


```
;
;
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-811-384-12

Query Match      21.5%; Score 756; DB 10; Length 450;
Best Local Similarity 28.6%; Pred. No 5e-26;
Matches 194; Conservative 79; Mismatches 138; Indels 268; Gaps 16;

QY 1 EVLOQSGDPLVKPGASVKISCKASYSTGYGMHWKSPGKLEWICRINPNNGVTLY 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 EVLVESGGGLVPGGSLRSCATSCYTETETMHWMRQAPGKLEWAGINPKNGGTSH 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 NQFKDKATLTVDKSSTAYMELRLSDESAVYYCARSTMITNY-----VMDYWGQGS 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 NQFMDFRTISVDKSTSTAYMQMNSLRAEDTAVYYCARMRGL-NYGFVRYFDVWGQGT 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 116 VTSSAKTPPSVYPLAGSAAQTNSMVLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPA 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 VTSSASTKGPSVFPAPLACRSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 176 VLQSD-LYTLSSSVTVPSSTWPSSETVTCNVAPASSTKVDKTIKVR-----DSG 223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 VLQSSGLYSLSSVTVTSSNFGQTTCNVDRHKPSNTKVDKTVVERKCCVECPAPPPVA 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 224 GPSEKEEINEKDLRKLSELOGTALGNLKOIYYNSKATISSEKSADQFNTLLPKGFF 283
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 GPS-----VFLFPKP-----KDTLMI----- 256

QY 284 TGHPMYNDLLVLGSLTAATSEYEGSSVDLYGAYGYQCAGGTENKTCACMYGVTLHDNNR 343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 -----SRPEVT-CVVVDVS----- 270

QY 344 LTEKKVPINLWTDGKOTTVPIDKVTSKKEVTVQBELDLQARHLHGKFLYNSDSFGGK 403
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 -HEDPEQFNWYDGM-----VHNAKTKPRE----- 296

QY 404 VQGLVIFHSSESTSYDLFDAGQGYPTDLLRIYRDNTTISSTLSLSLYLTTIVMT 463
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 -----EQNFSTRVSVLTVVHOD----- 315

QY 464 QTPTSLLSAGDRVTITCKASQSVNDVWYQOKPGSPKLLISYTSRYAGVPDRFSGS 523
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 -----WLNCKEYKC-----KVSNK 329

QY 524 GYGTDFTLTISVQAEAAVYFCQDYNSPPTFGGKLEIKRADAAPTVSIFFPSSEQL 583
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 GLPAPIEKTISKTKGQ-----PREPQVYTLPPSREEM 361

QY 584 TSCGASVCFNNFYPKDINVKKIDGSRQN-----GVNLNSWTDQDSKDSYMSSTLT 638
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 TKNQVSLTCLVKGFYPSDIAVENESGQPNYKTPPMLDS-----DGSEFLYSKLT 414

QY 639 LTKDEYERHNSYTCETHK 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 415 VDKSRWQGNVFCSCVNH 433

RESULT 28
US-09-799-514-8
; Sequence 8, Application US/09799514
; Patent No. US20020065220A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: Immunoglobulin Superfamily Polynucleotides, Polypeptides, and An
; FILE REFERENCE: PT015P1
; CURRENT APPLICATION NUMBER: US/09/799,514
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/US00/23662
```

```
;
;
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/152,248
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-514-8

Query Match      21.0%; Score 740; DB 10; Length 240;
Best Local Similarity 61.3%; Pred. No. 1.4e-25;
Matches 146; Conservative 32; Mismatches 50; Indels 10; Gaps 2;

QY 444 ISSTLSLSLYLVT-----SIVMTQPTSLLSVAGDRVTITCKASQSV-----SNDVAV 493
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 VLQTVFISLLLSWISGAYGDIVMTQSPDSLAVSLGERATINCKSSQTVLYSSNNKNYLA 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 494 YQOKPQSPKLLISYTSRYAGVPDRFSGSGYGTDTFTLTISVQAEAAVYFCQDYN 553
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 YQOKPQSPKLLIYWASTRESGVDPDRFSGSGCTDFTLTISLQAEADVAVYQCQYSTP 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 554 PTEGGTKLEIKRADAAPTVSIFFPSSEQLTSGASVCFNNFYPKDINVKKIDGSR 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 YSFGQTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVOMKVDNALQ 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 614 QNGVLNSWTDQDSKDSYMSSTLTLTCKEYERHNSYTCETHKSTSPVKSFN 671
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 SGNQSESVTEQDSKDSYMSSTLTLSKADYERKHYACEVTHQGLSSPVTKSFNRGE 239

RESULT 29
US-09-940-166A-6
; Sequence 6, Application US/09940166A
; Patent No. US20020058324A1
; GENERAL INFORMATION:
; APPLICANT: Blank, Gregory S.
; TITLE OF INVENTION: Protein Recovery
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,166A
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/097,309
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwartz, Timothy R.
; REGISTRATION NUMBER: 32171
; REFERENCE/DOCKET NUMBER: P1105R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-7467
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
```


Db 62 FSGSGTDTTITSSLOPEDATYVCOGNLTLPFTFGQGTVEIKRTVAAPSVFIFFPS 121
QY 580 SEQLTSGGASVVCFLNNFKPDINWKIDGSRQNGVLNSWTDODSKDSTYSMSLTLL 639
Db 122 DEQLKSGTASVCLLNNFVPRAKVQWVDNALSGNSQESVTEQDSKDSYLSLTLL 181
QY 640 TKDEYERHNSYTCETHKTSTSPVKSFRNE 671
Db 182 SKADYKHKVYACEVTHQGLSSPVTKSNRGE 213

RESULT 32
US-09-995-693-1
; Sequence 1, Application US/09995693
; Patent No. US20020136721A1
; GENERAL INFORMATION:
; APPLICANT: Schwall, Ralph H.
; Tabor, Kelly H.
; TITLE OF INVENTION: Hepatocyte Growth Factor Receptor
; Antagonists and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/995,693
; FILING DATE: 29-NO. US20020136721A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/952,235
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0938P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 220 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-693-1

Query Match 20.6%; Score 726; DB 10; Length 220;
Best Local Similarity 63.3%; Pred. No. 5,1e-25;
Matches 138; Conservative 34; Mismatches 40; Indels 6; Gaps 1;

QY 460 IYVMTPTLLYSAGRVITITCKASQSV-----SNDVAYQKPGQSPKLLISYTSRY 513
Db 2 IMWSQPSLLTVSVGEKVTVCSSQSLLYTSSQKNYLAWYQKPGQSPKLLIYNASTRE 61
QY 514 AGVPRFSSGSGTDFTLTISSVQAEADAVYFCQDYNPPTFGGKLEIKRADAAPTV 573
Db 62 SGVPRFTGSGSGTDFTLTITSVKADDLAVYVCQYIAYPWTFGGKLEIKRTVAAPSV 121
QY 574 SIFFPSSEQLTSGGASVVCFLNNFKPDINWKIDGSRQNGVLNSWTDQDSKDSYSM 633
Db 122 FIFPSDEQLKSGTASVCLLNNFVPRAKVQWVDNALSGNSQESVTEQDSKDSYSL 181
QY 634 SSTLTITKDEYERHNSYTCETHKTSTSPVKSFRNE 671

Db 182 SSTLTSLKADYKHKVYACEVTHQGLSSPVTKSNRGE 219

RESULT 33
US-09-940-166A-7
; Sequence 7, Application US/09940166A
; Patent No. US20020058324A1
; GENERAL INFORMATION:
; APPLICANT: Blank, Gregory S.
; Narindray, Daljit S.
; Zapata, Gerardo A.
; TITLE OF INVENTION: Protein Recovery
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,166A
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/097,309
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwartz, Timothy R.
; REGISTRATION NUMBER: 32171
; REFERENCE/DOCKET NUMBER: P1105R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-7467
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-940-166A-7

Query Match 20.6%; Score 725.5; DB 10; Length 300;
Best Local Similarity 53.5%; Pred. No. 7e-25;
Matches 147; Conservative 43; Mismatches 62; Indels 23; Gaps 5;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYSTGYIMHWKQSPGKGLIEWIGRINPNNGVTLY 60
Db 24 EVQLVESGGGLVQPGGSLRLSCATGYTETMTMMRWQAQPGKLEWVAGINPKNGGTS 83
QY 61 NQKFKDKATLTVDKSTTAYMELRSLTSDSAVYICARSTMTITNY-----VMDYWGQGS 115
Db 84 NORFMDRFTISVDKSTSTAYMOMNSLRADETAVYICARWRGL-NYGFVRYFDVWGQGL 142
QY 116 VTVSSAKTTPPSVYPLAPGSAAGTNSMTVLGGLVKGYPPEVTVTNWNSGLSSGVHTFPA 175
Db 143 VTVSSASTKGPSVFFLAPSSKSTSGTAALGCLVLDYDPEPVTVSNWNSGALTSGVHTFPA 202
QY 176 VLQSD-LYTLSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIVPRDS----- 222
Db 203 VLQSSGLYSLSSVTVPSSSLGTQTVICNVNHPKSNTKVDKKVEPKSCDKTHTCTPCPPAP 262
QY 223 ---GGPSEKSEETNEKDLRKKSELOGTALGNLKOI 254
Db 263 ELIGGRMKOLEDKVELLSKNYHLE-NEVARLKKL 296

```
RESULT 34
US-09-825-012-66
; Sequence 66, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 Fd - DNase I fusion
US-09-825-012-66

Query Match 20.5%; Score 722.5; DB 10; Length 515;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVQLQSGPDLVKPCASVKISKASGYSTGYMHVWVKQSPGKLEWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPGASVKVSKASGYTFSAYIEWVRQAPGKGLWYGEILPGSNNSRY 79
;
Qy 61 NQKFKDKATLVTKSSSTAYMELSLTSEDSAVYYCARSTMITNVMYDYGQGTSTVTVSS 120
Db 80 NEKFKGRVTVRTSTNTAYMELSLRSEDYAVYYCARSDYFAWFA--YWGQGTLLTVTVSS 137
;
Qy 121 AKTTPPSVYPLAPGSAAGTNSMTVLGCLVKGYFPEPTVTWNSGSLSSGVHTFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAAALGCLVKDYFPEPTVTSWNSGALTSGVHTFPAVLQSS 197
;
Qy 181 -LYTLSSSVTPSSWPSETVTCNVAHPASSTKVDKKIVPR 220
Db 198 GLYSLSVVTPSSSLGTQTYICNVNHHKPSNTKVDKKVEPK 238
;
RESULT 35
US-09-825-012-38
; Sequence 38, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 Fd - DNase I fusion
US-09-825-012-38

Query Match 20.5%; Score 722.5; DB 10; Length 517;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVQLQSGPDLVKPCASVKISKASGYSTGYMHVWVKQSPGKLEWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPGASVKVSKASGYTFSAYIEWVRQAPGKGLWYGEILPGSNNSRY 79
;
Qy 61 NQKFKDKATLVTKSSSTAYMELSLTSEDSAVYYCARSTMITNVMYDYGQGTSTVTVSS 120
Db 80 NEKFKGRVTVRTSTNTAYMELSLRSEDYAVYYCARSDYFAWFA--YWGQGTLLTVTVSS 137
;
Qy 121 AKTTPPSVYPLAPGSAAGTNSMTVLGCLVKGYFPEPTVTWNSGSLSSGVHTFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAAALGCLVKDYFPEPTVTSWNSGALTSGVHTFPAVLQSS 197
;
Qy 181 -LYTLSSSVTPSSWPSETVTCNVAHPASSTKVDKKIVPR 220
Db 198 GLYSLSVVTPSSSLGTQTYICNVNHHKPSNTKVDKKVEPK 238
;
RESULT 36
US-09-825-012-76
; Sequence 76, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 Fd - DNase I fusion
US-09-825-012-76

Query Match 20.5%; Score 722.5; DB 10; Length 519;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVQLQSGPDLVKPCASVKISKASGYSTGYMHVWVKQSPGKLEWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPGASVKVSKASGYTFSAYIEWVRQAPGKGLWYGEILPGSNNSRY 79
;
Qy 61 NQKFKDKATLVTKSSSTAYMELSLTSEDSAVYYCARSTMITNVMYDYGQGTSTVTVSS 120
Db 80 NEKFKGRVTVRTSTNTAYMELSLRSEDYAVYYCARSDYFAWFA--YWGQGTLLTVTVSS 137
;
Qy 121 AKTTPPSVYPLAPGSAAGTNSMTVLGCLVKGYFPEPTVTWNSGSLSSGVHTFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAAALGCLVKDYFPEPTVTSWNSGALTSGVHTFPAVLQSS 197
;
Qy 181 -LYTLSSSVTPSSWPSETVTCNVAHPASSTKVDKKIVPR 220
Db 198 GLYSLSVVTPSSSLGTQTYICNVNHHKPSNTKVDKKVEPK 238
;
RESULT 37
US-09-825-012-80
; Sequence 80, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
```

; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFGL Fd - DNase I fusion
US-09-825-012-80

Query Match 20.5%; Score 722.5; DB 10; Length 519;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVLOQSGDPLVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGRIINPNNGVTLY 60
:||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Db 20 QVQLVQSGAEVKKPGASVKVSCKASGYTFSAYIEWVRQAPGKLEWGEILPGSNNSRY 79
:||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Qy 61 NQKFKDKATLTVDKSTTAYMELRLSITSDSAVYYCARSTMITNTYVMDYWGOGTSVTYSS 120
:||| : ||: ||: ||||| || ||||| || ||||| : ||||| |||||
Db 80 NEKFGKRVTVTRDTSTNTAYMELSLRSEDYAVYYCARSYDFAWFA--YWGOGTLVTYSS 137
:||| : ||: ||: ||||| || ||||| || ||||| : ||||| |||||
Qy 121 AKTTPPSVYPLAPGSAAGTNSMVTLCGLVKGYFPEPVTVTWNSGSLSGVHTFPVAVLQSD 180
: || ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 ASTKGPSVPEPLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPVAVLQSS 197
: || ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 -LYTLSSSVTPSPSTWPTSETVTCNVAHPASSTKVDKKIVPR 220
:||||| ||||| : || ||| || |||||: ||: |||||: ||:
Db 198 GLYSLSVVTPSSSLGTQTYICNVNHNKPSNTKVDKKVEPK 238
:||||| ||||| : || ||| || |||||: ||: |||||: ||:

RESULT 38
US-09-825-012-71
; Sequence 71, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFGL Fd - DNase I fusion
US-09-825-012-71

Query Match 20.5%; Score 722.5; DB 10; Length 521;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVLOQSGDPLVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGRIINPNNGVTLY 60
:||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Db 20 QVQLVQSGAEVKKPGASVKVSCKASGYTFSAYIEWVRQAPGKLEWGEILPGSNNSRY 79
:||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Qy 61 NQKFKDKATLTVDKSTTAYMELRLSITSDSAVYYCARSTMITNTYVMDYWGOGTSVTYSS 120
:||| : ||: ||: ||||| || ||||| || ||||| : ||||| |||||
Db 80 NEKFGKRVTVTRDTSTNTAYMELSLRSEDYAVYYCARSYDFAWFA--YWGOGTLVTYSS 137
:||| : ||: ||: ||||| || ||||| || ||||| : ||||| |||||
Qy 121 AKTTPPSVYPLAPGSAAGTNSMVTLCGLVKGYFPEPVTVTWNSGSLSGVHTFPVAVLQSD 180
: || ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 ASTKGPSVPEPLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPVAVLQSS 197
: || ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 181 -LYTLSSSVTPSPSTWPTSETVTCNVAHPASSTKVDKKIVPR 220
:||||| ||||| : || ||| || |||||: ||: |||||: ||:
Db 198 GLYSLSVVTPSSSLGTQTYICNVNHNKPSNTKVDKKVEPK 238
:||||| ||||| : || ||| || |||||: ||: |||||: ||:

RESULT 39
US-09-825-012-85
; Sequence 85, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 85
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFGL Fd - DNase I fusion
US-09-825-012-85

Query Match 20.5%; Score 722.5; DB 10; Length 525;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVLOQSGDPLVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGRIINPNNGVTLY 60
:||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Db 20 QVQLVQSGAEVKKPGASVKVSCKASGYTFSAYIEWVRQAPGKLEWGEILPGSNNSRY 79
:||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Qy 61 NQKFKDKATLTVDKSTTAYMELRLSITSDSAVYYCARSTMITNTYVMDYWGOGTSVTYSS 120
:||| : ||: ||: ||||| || ||||| || ||||| : ||||| |||||
Db 80 NEKFGKRVTVTRDTSTNTAYMELSLRSEDYAVYYCARSYDFAWFA--YWGOGTLVTYSS 137
:||| : ||: ||: ||||| || ||||| || ||||| : ||||| |||||
Qy 121 AKTTPPSVYPLAPGSAAGTNSMVTLCGLVKGYFPEPVTVTWNSGSLSGVHTFPVAVLQSD 180
: || ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 ASTKGPSVPEPLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPVAVLQSS 197
: || ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 -LYTLSSSVTPSPSTWPTSETVTCNVAHPASSTKVDKKIVPR 220
:||||| ||||| : || ||| || |||||: ||: |||||: ||:
Db 198 GLYSLSVVTPSSSLGTQTYICNVNHNKPSNTKVDKKVEPK 238
:||||| ||||| : || ||| || |||||: ||: |||||: ||:

RESULT 40
US-09-825-012-43
; Sequence 43, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFGL Fd - DNase I fusion

ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: PL729C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-5530
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 232 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-811-384-10

Query Match 20.4%; Score 718.5; DB 10; Length 232;
Best Local Similarity 61.1%; Pred. No. 1.1e-24;
Matches 138; Conservative 35; Mismatches 46; Indels 7; Gaps 3;
QY 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYMHVWQSPGKGLGWIGRINPNNGVTLY 60
DB 1 EVQLVESGGGLVQPGGSLRLSCATSGYTFETETVTHWMRQAPGKGLGWAGINPKNGGTS 60
QY 61 NQKFKDKATLTVDKSSITAYMELRSLTSEDSAVYYCARSTMTNY-----VMDYWGQGTS 115
DB 61 NQRFMDRFTISVDKSTSTAYMQMNSLRAEDTAVYYCARWRGL-NYGFDRVYFDVWGQGL 119
QY 116 VTSSAKTTPPSVYPLAPGSAQTNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPA 175
DB 120 VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA 179
QY 176 VLQSD-LYTLSSSVTPSPSTWPTETVCNVAHPASSTKVDKKIIVPR 220
DB 180 VLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHPKPSNTKVDKKVEPK 225

RESULT 44
US-09-940-166A-1
Sequence 1, Application US/09940166A
Patent No. US2002058324A1
GENERAL INFORMATION:
APPLICANT: Blank, Gregory S.
Narindray, Daljit S.
Zapata, Gerardo A.
TITLE OF INVENTION: Protein Recovery
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/940,166A
FILING DATE: 27-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/097,309
FILING DATE: 13-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Schwartz, Timothy R.
REGISTRATION NUMBER: 32171
REFERENCE/DOCKET NUMBER: PL105R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-7467
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 241 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-940-166A-1
Query Match 20.4%; Score 718.5; DB 10; Length 241;
Best Local Similarity 61.1%; Pred. No. 1.1e-24;
Matches 138; Conservative 35; Mismatches 46; Indels 7; Gaps 3;

QY 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYMHVWQSPGKGLGWIGRINPNNGVTLY 60
DB 1 EVQLVESGGGLVQPGGSLRLSCATSGYTFETETVTHWMRQAPGKGLGWAGINPKNGGTS 60
QY 61 NQKFKDKATLTVDKSSITAYMELRSLTSEDSAVYYCARSTMTNY-----VMDYWGQGTS 115
DB 61 NQRFMDRFTISVDKSTSTAYMQMNSLRAEDTAVYYCARWRGL-NYGFDRVYFDVWGQGL 119
QY 116 VTSSAKTTPPSVYPLAPGSAQTNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPA 175
DB 120 VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA 179
QY 176 VLQSD-LYTLSSSVTPSPSTWPTETVCNVAHPASSTKVDKKIIVPR 220
DB 180 VLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHPKPSNTKVDKKVEPK 225

RESULT 45
US-09-736-371B-21
Sequence 21, Application US/09736371B
Patent No. US20020131968A1
GENERAL INFORMATION:
APPLICANT: Waldmann, Herman
APPLICANT: Frewin, Mark
TITLE OF INVENTION: AGLYCOSYLATED ANTIBODIES
FILE REFERENCE: Waldmann
CURRENT APPLICATION NUMBER: US/09/736,371B
CURRENT FILING DATE: 2002-04-25
PRIOR APPLICATION NUMBER: 9815909.8
PRIOR FILING DATE: 1998-07-21
PRIOR APPLICATION NUMBER: PCT/GB99/02380
PRIOR FILING DATE: 1999-07-21
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 449
TYPE: PRT
ORGANISM: Homo sapiens
US-09-736-371B-21

Query Match 20.4%; Score 718.5; DB 10; Length 449;
Best Local Similarity 29.3%; Pred. No. 1.9e-24;
Matches 194; Conservative 71; Mismatches 161; Indels 237; Gaps 17;
QY 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYMHVWQSPGKGLGWIGRINPNNGVTLY 60
DB 1 EVQLLESGLVQPGGSLRLSCAASGFTFSFPAWVQAPGKGLGWYSTISTSGRTYY 60
QY 61 NQKFKDKATLTVDKSSITAYMELRSLTSEDSAVYYCARSTMTNYVMDYWGQGTSTVTVSS 120
DB 61 RDSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKFRQYSG-GFDYWGQGLTVTVSS 119
QY 121 AKTTPPSVYPLAPGSAQTNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
DB 120 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS 179
QY 181 -LYTLSSSVTPSPSTWPTETVCNVAHPASSTKVDKKIIVPRDSGGSPSEKSEINEKDLRK 239
DB 180 GLYSLSSVTVTPSSSLGTQTYICNVNHPKPSNTKVDKKVEPKSC----- 222
QY 240 KSELQGTALGNLKOIYYNYSKAITSSSEKSDQFLTNTLLFKGFTGHPWYNLLVLQGST 299

Db 223 -----DKTHTPCPCAPELIGGPSVF--LFPKPK--KDTLM----- 254
Qy 300 AATSEYEGSSVDLYGAYGQCAGGTGPNKTACMYGGVTLHDNNRLTEKKVPINLWDGK 359
Db 255 -----ISRPEVT-CVVVDVS-----HEDPEVKFNWYDVG- 283
Qy 360 QTTVPIDKVKTSKKEVTAVQELDLQARHYLHGKFGLYNSDFGKGVQRLIVFHSSEGSTV 419
Db 284 --VEVIHNAKTKPRE----- 295
Qy 420 SYDLFDAQGQVPDPTLLRIYRDNNTTISSTLSISLYLTTTSIVMTQPTTSLLSVAGDRVTI 479
Db 296 -----EQYAST-YRWVSVLTVLHODWLNKEY----- 321
Qy 480 TKASQSVSNDVAVYQOKPGOSPKLLISYTSRSGVAGVPDRFSGSGYCTDFTLTISVQAE 539
Db 322 CKKVS-----NKALPAP---TEKTISKAKGP----- 345
Qy 540 DAAVYFCQDYNSPPTFGGKTLEIKRADAAPTVSIPIPPSSEQLTSGGASVVCFLNNFYP 599
Db 346 -----REPQYITLPPSRDELTKNOVSLTCLVKGFYP 376
Qy 600 KQINVKWIDGSEKQ-----GVLSNWDQDQSKDSTYSMSSTLTLTAKDEYERHNSYTCEA 654
Db 377 SDIAVENESNGQPNKYKTPPVLDSD-----DGSFELYSKLTVDKSRWQOGNVFSCSV 429
Qy 655 THK 657
Db 430 MHE 432

RESULT 46
US-09-828-995B-17
; Sequence 17, Application US/09828995B
; Patent No. US20020165135A1
; GENERAL INFORMATION:
; APPLICANT: Heskia Corporation
; APPLICANT: McCall, Catherine A.
; APPLICANT: Tang, Liang A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATED TO CANINE IgG AND CANINE IL-13
; FILE REFERENCE: AL-7
; CURRENT APPLICATION NUMBER: US/09/828,995B
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,874
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/195,659
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Canis familiaris
; US-09-828-995B-17

Query Match 20.4%; Score 717.5; DB 9; Length 474;
Best Local Similarity 27.4%; Pred. No. 2.2e-24;
Matches 186; Conservative 80; Mismatches 149; Indels 265; Gaps 13;
Qy 1 EVQLQQSGDPLVKPGASVKISKASGYSFTGYMHWVKQSPGKLEWIGRIINPNNGVTLY 60
Db 20 DVQLVESGDDLKPGGSLRLSCVASGFTFSSCAMSWVRQSPGKQPWVATIRYDGSIIY 79
Qy 61 NQKFKDRLTVKSSITAYMELRSLTSEDSAVYICARSTMIT--NYVMDVWGGQTSVTV 118
Db 80 ADAYKGFSTRDAKNTVYLQMSLRAEDTAVYICAKAPPDYSHYGMVWGPSTSLFV 139
Qy 119 SSAKTTPEVYPLAPGSAATQMSMTVLGCLVKGVFPPEVTVTNWNSGLSSGVHTEFPAVLQ 178
Db 140 SSASTTAPVFPPLAPSGCGSGSTVALACLVSIGYIPEVTVSNWSVLTSGVHTFSPVLQ 199
Qy 179 SD-LYTLSSSVYVPSSTWPSSTVTCNVHPASSTKVDKIKVPRDS----- 222

Db 200 SSGLYSLSMVTVPSRRWPSSETFTCNVAHPATNTKVDKPVAKCECEKCNMCNCPGCGL 259
Qy 223 -GGPSEKSEELNEKDLRKKSELOQTALGNLQKIYYYSKAITSEKESADQPLTWTLLPKG 281
Db 260 LGGPS----- 264
Qy 282 FFTGHPWYNLLVDLSTAATAATSEYEGSSVDLYGAYGQCAGGTGPNKTACMYGGVTLHDN 341
Db 265 VFIFPPKPKDLV---TARTPTVTCVVVDLD----- 292
Qy 342 NRUTEKKVPINLWDGKQ-----TTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNS 397
Db 293 ---PENPEVQISWFDVSKQVQTANTQPREQSNQTYRV-VSVLPIGHQDWLSGK----- 342
Qy 398 DSGGKQVQRLIVFHSSEGSTVSVDLFDAGQOYPTDRLRIYRDNNTTISSTLSISLYLT 457
Db 343 -QFKCKV-----NNKALPS----- 355
Qy 458 TSIVMTQPTTSLLSVAGDRVTITCKASQSVSNDVAVYQOKPGOSPKLLISYTSRSGVAGVP 517
Db 356 -----PIEELIS-----KTPGOAHQ----- 370
Qy 518 DRFSGSGYGTDFTLTISVQAEAAVYFCQDYNSPPTFGGKTLEIKRADAAPTVSIFF 577
Db 371 -----PNVYVLP 377
Qy 578 PSSEQLTSGGASVVCFLNNFYPKIDINVKWKIDGSEKQNGVLSNWDQDQSKDSTYSMSSTL 637
Db 378 PSRDEMSKNTVTLTCLVKDFPPEIDVEMQSGQOEPESKYRMTPPOLDEGVSFLYSKL 437
Qy 638 TLTAKDEYERHNSYTCEATHK 657
Db 438 SVDKSRWQRGDTFICAVMHE 457
RESULT 47
US-10-011-125-5
; Sequence 5, Application US/10011125
; Patent No. US20020142388A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Christina Yu-Ching
; TITLE OF INVENTION: BACTERIAL HOST STRAINS
; FILE REFERENCE: F1804R1
; CURRENT APPLICATION NUMBER: US/10/011,125
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/256,162
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 5
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence is synthesized.
US-10-011-125-5

Query Match 20.3%; Score 716; DB 12; Length 212;
Best Local Similarity 64.2%; Pred. No. 1.3e-24;
Matches 136; Conservative 29; Mismatches 45; Indels 2; Gaps 1;
Qy 460 IVMTQPTTSLLSVAGDRVTITCKASQSVSNDVAVYQOKPGOSPKLLISYTSRSGVAGVPDR 519
Db 2 IQMTQSPSSLSASGDRVTITCRASQDINNLNYYQKPGKAPKLLIYTS--HSGVPSR 59
Qy 520 FSGSGYGTDFTLTISVQAEAAVYFCQDYNSPPTFGGKTLEIKRADAAPTVSIFFPS 579
Db 60 FSGSGSGTDTLTISLQPEDFATYYCQGNLTLPPTFGGKTLEIKRTVAAPSVFIFFPS 119
Qy 580 SEQLTSGGASVVCFLNNFYPKIDINVKWKIDGSEKQNGVLSNWDQDQSKDSTYSMSSTLT 639
Db 120 DEQLKSGTASVVCFLNNFYPKIDINVKWKIDGSEKQNGVLSNWDQDQSKDSTYSMSSTLT 179
Qy 640 TKDEYERHNSYTCEATHKSTSPIVKSFNRNE 671

Db 180 SKADYEHKKYACEVTHQGLSSPVTKSFNRGE 211

RESULT 48

US-10-006-771A-6

; Sequence 6, Application US/10006771A

; Patent No. US20020165360A1

; GENERAL INFORMATION:

; APPLICANT: Junghans, Richard P.

; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen

; FILE REFERENCE: 002

; CURRENT APPLICATION NUMBER: US/10/006,771A

; CURRENT FILING DATE: 2002-06-04

; PRIOR APPLICATION NUMBER: 60/250,090

; PRIOR FILING DATE: 2000-11-30

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 232

; TYPE: PRT

; ORGANISM: Homo sapiens and Mus sp.

US-10-006-771A-6

Query Match 20.3%; Score 715.5; DB 9; Length 232;

Best Local Similarity 61.7%; Pred. No. 1.5e-24;

Matches 140; Conservative 30; Mismatches 50; Indels 7; Gaps 2;

QY 451 ISLYLYTTS-----IWTQTPTSLVYAGDRVTITCKASQSVNSDVAVYQQRPGQSPKL 504

Db 6 IILFLVATATGVHSIDQLTQSPSLASVGDRTVITCKASQDVGTSVAVYQQRFGKAPKL 65

QY 505 LISYTSRYAGVDPDRFSGSGYGTDTLTISVQAEADAAYFCQODYNSPPTFFGGGTKLEI 564

Db 66 LIYVTSRHTGVSPRFSGSGGTDTFTTISLQPEDIATYYCQ-YSLYRSFGGQTKVEI 124

QY 565 KRADAAPTIVSIFPPSSQLTSGGASVCFVLFNNFYPKDINVKWKIDGSEKNGVLSNWTQD 624

Db 125 KRTVAAPSVFIFPPSDQLKSGTASVYCLINNFYPREAKVQWKVDNALQSGNSQESVTEQ 184

QY 625 DSKDSTVSMSTLTLTKDEYERHNSYTCETHKTSTSTPIVKSFNRRNE 671

Db 185 DSKDSTVLSLSTLTLSKADYEHKKYACEVTHQGLSSPVTKSFNRGE 231

RESULT 49

US-09-249-011A-22

; Sequence 22, Application US/09249011A

; Patent No. US20020176855A1

; GENERAL INFORMATION:

; APPLICANT: CO, MAN SUNG

; APPLICANT: VASQUEZ, MAXIMILIANO

; APPLICANT: CARENNO, BEATRIZ

; APPLICANT: CELNIKER, ABBIE CHERYL

; APPLICANT: COLLINS, MARY

; APPLICANT: GOLDMAN, SAMUEL

; APPLICANT: GRAY, GARY S.

; APPLICANT: KNIGHT, ANDREA

; APPLICANT: O'HARA, DENISE

; APPLICANT: RUP, BONITA

; APPLICANT: VELDMAN, GEERTRUIDA M.

; TITLE OF INVENTION: HUMANIZED IMMUNOGLOBULIN REACTIVE WITH B7-2 AND METHODS

; FILE REFERENCE: 08702.0081-000000

; CURRENT APPLICATION NUMBER: US/09/249, 011A

; CURRENT FILING DATE: 1999-02-12

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 22

; LENGTH: 239

; TYPE: PRT

; ORGANISM: Mus sp.

US-09-249-011A-22

Search completed: January 8, 2003, 12:05:26
Job time : 25.6232 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:55:11 ; Search time 25.2464 Seconds
(without alignments)
2558.872 Million cell updates/sec

Title: US-09-900-766-1
Perfect score: 3522
Sequence: 1 EVQLQQSGDLVKPEASVKI.....EATHKTSTPIVKSFNRNES 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1107	31.4	257	A28179	enterotoxin E prec
2	1036.5	29.4	225	S37484	Ig kappa chain - m
3	973	27.6	446	A40295	Ig gamma-2a chain
4	952	27.0	220	A31790	Ig kappa chain V r
5	951.5	27.0	469	S37483	Ig gamma-2a chain
6	948	26.9	257	A28664	enterotoxin A prec
7	931	26.4	548	S38864	Ig epsilon chain C
8	929	26.4	246	S38950	Ig gamma chain - m
9	926	26.3	214	S68212	Ig kappa chain (Ma
10	918	26.1	260	C83984	enterotoxin P [imp
11	914.5	26.0	219	S52028	Ig kappa chain - m
12	914	26.0	214	PC4202	monoclonal antibod
13	910.5	25.9	219	PC4203	Ig kappa chain (mo
14	903.5	25.7	219	S16112	Ig kappa chain V r
15	900.5	25.6	474	G2MS11	Ig gamma-2b chain
16	896.5	25.5	219	S38865	Ig kappa chain - m
17	892	25.3	218	JC5810	monoclonal antibod
18	890.5	25.3	221	S49220	Ig gamma-1 chain -
19	886.5	25.2	217	S42772	Ig kappa chain - m
20	880	25.0	235	S68241	Ig kappa chain V r
21	878.5	24.9	235	S25058	Ig kappa chain - m
22	876.5	24.9	225	JL0029	Ig kappa chain pre
23	875	24.8	234	S14237	Ig kappa chain pre
24	874	24.8	234	S01320	Ig kappa chain pre
25	873	24.8	475	S01321	Ig gamma-2b chain
26	870	24.7	444	PC4436	monoclonal antibod
27	857	24.3	210	A56169	Ig kappa chain V r
28	853	24.2	240	S06084	Ig kappa chain pre
29	808	22.9	220	S68211	Ig heavy chain (Ma

ALIGNMENTS

RESULT 1

A28179 enterotoxin E precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 15-Oct-1999
C:Accession: A28179
R:Couch, J.L.; Soltis, M.T.; Betley, M.J.
J. Bacteriol. 170, 2954-2960, 1988
A:Title: Cloning and nucleotide sequence of the type E staphylococcal enterotoxin gen
A:Reference number: A28179; MUID:88257005; PMID:3384800
A:Accession: A28179
A:Molecule type: DNA
A:Residues: 1-257 <COU>
A:Cross-references: GB:M21319; MID:gl53001; PIDN:AAA26617.1; PID:gl53002
C:Superfamily: enterotoxin B

Query Match 31.4%; Score 1107; DB 2; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-50;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SKSESEINKEKLRKSELGCTALGNLKKQIYYNNSKAITSEKSAKQDPLNTLLFKGFFTG 285
|||||
Db 25 SKSESEINKEKLRKSELGCTALGNLKKQIYYNNSKAITSEKSAKQDPLNTLLFKGFFTG 84
QY 286 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMGVTLHDNNRLT 345
|||||
Db 85 HPWYNDLLVDLGSKDATNKYKGVLDLYGAYGYQCAGGTPNKTACMGVTLHDNNRLT 144
QY 346 EEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKQV 405
|||||
Db 145 EEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKQV 204
QY 406 RGLIVFHSSEGSTSVSYDLFDAQGQVPTDLLRIYRDNTTISSTLSLSLYLT 458
|||||
Db 205 RGLIVFHSSEGSTSVSYDLFDAQGQVPTDLLRIYRDNTTISSTLSLSLYLT 257

RESULT 2

S37484 Ig kappa chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
C:Accession: S37484
R:Ducancel, F.F.D.
submitted to the EMBL Data Library, February 1993
A:Reference number: S37483
A:Accession: S37484
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-225 <DUC>
A:Cross-references: EMBL:X70424; NID:g406254; PIDN:CAA49869.1; PID:g406255
C:Superfamily: immunoglobulin V region; immunoglobulin homology

DB 182 SSTLTTRDEYERHNSYTCEATHKTSPIVKSFNENE 219

DB 182 SSTLTTRDEYERHNSYTCEATHKTSPIVKSFNRE 219

RESULT 5

S37483
Ig gamma-2a chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37483
R:Ducancel, F.F.D.
submitted to the EMBL Data Library, February 1993
A:Reference number: S37483
A:Accession: S37483
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-469 <DUC>
A:Cross-references: EMBL:X70423; NID:g406252; PIDN:CAA9868.1; PID:g406253
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:276-345/Domain: immunoglobulin homology <IMM>

Query Match 27.0%; Score 951.5; DB 2; Length 469;
Best Local Similarity 34.2%; Pred. No. 3.6e-42;
Matches 236; Conservative 62; Mismatches 126; Indels 267; Gaps 13;
QY 1 EVOLQSGPDLVKPGASVKISKASGYSTGYVMHWKSPGKGLIEWIGRINPNNGVTLY 60
DB 20 QIOLQSGPDLVKPGASVKISKASGYSTGYVMHWKSPGKGLIEWIGRINPNNGVTLY 79
QY 61 NOKFKDKATLTVDKSSSTAYMELRLSITSEDSAVYYCARSTMITNYYMDYWGQTSVTSS 120
DB 80 NENFKCKATLTVDTSSTAYMQLSLSITSEDTAVYFCARAMGATATLLDYGQGTTLTVSS 139
QY 121 AKTTPPSVYPLAPGSAQNSMYTLGCLVKGYFPEPVTVTWNSGSLSSGVHFPFAVLQSD 180
DB 140 AKTTPSVYPLAPVCGDTTGTSSVTLGCLVKGYFPEPVTVTWNSGSLSSGVHFPFAVLQSD 199
QY 181 LYTLSSTVYVPSSTWPSSEITCNVHPASSTKVDKIKVPRDSGGPSEKSEINEKDLRKK 240
DB 200 LYTLSSTVYVPSSTWPSSEITCNVHPASSTKVDKIKIEPR-----GPIKPCP----- 247
QY 241 SELQGTALGNLKOIYYYNKAITSEKSADQFLNTLLFKGFTTGHWPYNLDLLVLDGSTA 300
DB 248 ----- 247

QY 301 ATSEYEGSSVDLYGAYGYQCAGTPNKTACMYGGVTLHDNNRLTEKKVPINLWDGKQ 360
DB 248 -----PKCKPAPN-----LLGGPSVF----- 263
QY 361 TTVPIDKVKTSKEVTVQELDLQARHYLHGKGLYNSDSFGGKVORGLIVFHSSEGSTVS 420
DB 264 -----IFPPKIKDVLMI-----SLSPVIT 282
QY 421 YDLFDAQGYQPTLLRIYRDNTTISSTLSISLYLTTISIVMTQTPTSLLSAGDRVTIT 480
DB 283 CVVVDVSEDDPD-----VQISWVFNVEVHTAQTOT----- 313
QY 481 CKASQSVNDVAVYQKPGQPKGLISYTSRYAGVPDRFSGSGYGTDTTLTSSVQAEQ 540
DB 314 -----HREDYNSTLRVY----- 325

QY 541 AAVYFCQDYNPPPTFGGKTLEIKRAD-----AAPTYSIFPPSSEQLT 584
DB 326 SALPIQHQQWMSGKER-----KCKVNNKDLPIERTISPKGSVRAPQVIVLPPEEEMT 381
QY 585 SGASVYCFNLNFPYKPDINVKWKIDGSRQN-----GVLSNMTDQDSKDSYMSMTSLTL 639
DB 382 KQVTLTCVMTDFMPEDIYVEVTNNKTELNYKNTPEVLDS-----DGSYFMYSKLRV 434
QY 640 TKDEYERHNSYCEATHK-TSTSPIVKFSNR 669
DB 435 EKKNNWERNYSYSCVSVHEGLHNNHHTTKFSR 465

RESULT 6

A28664

enterotoxin A precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 15-Oct-1999
C:Accession: A28664; A29566
R:Betley, M.J.; Mekalanos, J.J.
J. Bacteriol. 170, 34-41, 1988
A:Title: Nucleotide sequence of the type A staphylococcal enterotoxin gene.
A:Reference number: A28664; MUID:88086892; PMID:3335483
A:Accession: A28664
A:Molecule type: DNA
A:Residues: 1-257 <BET>
A:Cross-references: GB:M18970; NID:g153120; PIDN:AAA26681.1; PID:g153121
A:Experimental source: strain FRI337
R:Huang, I.Y.; Hughes, J.L.; Bergdoll, M.S.; Schantz, E.J.
J. Biol. Chem. 262, 7006-7013, 1987
A:Title: Complete amino acid sequence of staphylococcal enterotoxin A.
A:Reference number: A29566; MUID:87222293; PMID:3584106
A:Accession: A29566
A:Molecule type: protein
A:Residues: 25-241, 'S', 243-257 <HUA>
C:Genetics:
A:Gene: entA
A:Map position: 6
C:Superfamily: enterotoxin B

Query Match 26.9%; Score 948; DB 2; Length 257;
Best Local Similarity 76.4%; Pred. No. 2.6e-42;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;
QY 226 SEKSEINEKDLRKSELOQTALGNLKOIYYYNKAITSEKSADQFLNTLLFKGFTTG 285
DB 25 SEKSEINEKDLRKSELOQTALGNLKOIYYYNKAITSEKSADQFLNTLLFKGFTTD 84
QY 286 HPWYNLDLLVDLGAATASTSEGVSSVDLYGAYGYQCAGTPNKTACMYGGVTLHDNNRLT 345
DB 85 HPWYNLDLLVDLGAATASTSEGVSSVDLYGAYGYQCAGTPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKVPINLWDGKQNTVPLETVKTNKNTVQELDLQARHYLHGKGLYNSDSFGGKVQ 405
DB 145 EEKVPINLWDGKQNTVPLETVKTNKNTVQELDLQARHYLHGKGLYNSDSFGGKVQ 204
QY 406 RGLIVFHSSEGSTVSVDLFDAGQGYQPTLLRIYRDNTTISSTLSISLYLTT 458
DB 205 RGLIVFSTEPSVNYDLFGAGQGYQNTLLRIYRDNKTINSENMHIDIYLYTS 257

RESULT 7

S38864
Ig epsilon chain C region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001
C:Accession: S38864
R:Kipp, B.; Becker, W.; Schlaak, M.
submitted to the EMBL Data Library, November 1993
A:Description: Combination of a defined specificity and desired isotype by cloning of
A:Reference number: S38864
A:Accession: S38864
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-548 <KIP>
A:Cross-references: EMBL:227397; NID:g416537; PIDN:CAA81788.1; PID:g940782
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:353-421/Domain: immunoglobulin homology <IMM>

Query Match 26.4%; Score 931; DB 2; Length 548;
Best Local Similarity 35.0%; Pred. No. 5e-41;
Matches 241; Conservative 83; Mismatches 186; Indels 178; Gaps 19;
QY 1 EVOLQSGPDLVKPGASVKISKASGYSTGYVMHWKSPGKGLIEWIGRINPNNGVTLY 60
DB 1 QVKLLESGGDLVKPGGSLKLSAASGLTFTSSYGMKSVRQIPDKRLEWATISSGGTYTY 60
QY 61 NOKFKDKATLTVDKSSSTAYMELRLSITSEDSAVYYCARSTMITNYYMDYWGQTSVTSS 120

```
Db 61 PDVSKGRFTISRDNKNTLYLQWSSLSKSDTAMCYCARQGVSTMIREFAYWGQTLTVSA 120
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Qy 121 AKTTPSVPLAPGSAOATNSWTLGCLVKGYPPEVTVTWNSGSLSSGVHFFPAVLQSD 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 GKTTPSVPLAPGSAOATNSWTLGCLVKGYPPEVTVTWNSGSLSSGVHFFPAVLQSD 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 LYTSSSVTPSPSTWSEVTCNVAPASSTKVDKKIVPRDSG-----GPSEKSEI--- 232
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 LYTSSSVTPSPSTWSEVTCNVAPASSTKVDKKIVPRDSGCKPCIVPEVSSVFIFPP 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 233 NEKDLRKKSELQ-----GTALGNLKOIYYNSKAITSEKSDAQDPLTNTLLFKGFETH 286
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 241 KPDVLRSTRITQIYCFIYGHILNDVSVWMLDDREITDT-----LAQTVLIK----- 287
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Qy 287 PWYNDLLVDLGSTAAATSEGGSDVLYGAYGYOCAGGTPNKTACWYGGVTLHDNNRLTE 346
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 288 -----EEGKLASTCSKLNITEQQWMSSEFTCK-----VTSQGVLYLAH 326
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Qy 347 EKKVPLNLTIDGKQTV-----PIDKVKTSKKEVTQVELDLOARHYLHCKFGLYNSD 398
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 327 TRCP-----DHEPRGVITYLPPSLDLYQNGAPKLTCLVDLESEKNVN----- 372
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Qy 399 SFGGKQVRLIVFHSSEGSTVSDLEDAQOQVDPDTLRIYRONTTISSTLSISLYLTT 458
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 373 -----VTNQEKKTSVS-----ASQWYTK-----HNNNATTSITSI----- 403
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Qy 459 SIWMTQTPPSLLYSAGDRVTITCKASQSVNDVAVWYQKPGQSPKLLISYTSRYAGVDP 518
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 404 -----LPVAKD-----WIE----- 413
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Qy 519 RFGSGYGTDFTLTISVQAEDAAYFCODYNS--PPTFGGKLEIKRADAAPTVSIF 576
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 414 -----CYG-----YQCIVDHPDPKPIVRSITKPGQR--SAPVYVF 449
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Qy 577 PPSSEQLTGGASVWCFLLNFPKIDNVKWKIDGSRQNGVLSWTDQDSKDS--YSMS 634
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 450 PPPEEE--SEDKRTLCLIONFFPEDISVQWLGDKLISNSQHSSTTTPLKSNNGSNRGGFFIF 508
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Qy 635 SFTLTLDKDEYERHNSYTCEATHKTSTSP 662
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 509 SRLEVAKTLWTORKOFTQCVIHEALQKP 536
      || : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

RESULT 8
Ig gamma chain - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S38950
R:Klebert, S.; Kratzin, H.D.; Zimmermann, B.; Vaesen, M.; Frosch, M.; Weisgerber, C.; Bi
Biol. Chem. Hoppe-Seyler 374, 993-1000, 1993
A:Title: Primary structure of the murine monoclonal IgG2a antibody mAb735 against alpha
A:Reference number: S38950; MUID:94128242; PMID:8297501
A:Accession: S38950
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-246 <KLE>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:137-201/Domain: immunoglobulin homology <IM>

Query Match 26.4%; Score 929; DB 2; Length 246;
Best Local Similarity 78.5%; Pred. No. 2.3e-41;
Matches 179; Conservative 18; Mismatches 25; Indels 6; Gaps 2;

Qy 1 EVQLQSGDPLVKPGASVKISKASGYSTGYTHHWKQSPGKLEWIGRIINPENGVTLY 60
      : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 1 QIQLOQSGPELVRPGASVKISKASGYTFTDYIHWKQRPGEGLWIGWIYPGSGNTKY 60
      : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |

Qy 61 NQKFKDKATLTVDKSTTAYMELRSLTSEDSAVYCYCARSTMTITNYVMYWGQGTSTVSS 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 NEKFKGKATLTVDTSSTAYMOLSLTSEDSAVYFCARG---GKFAMDYWGQGTSTVSS 117
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Qy 121 AKTTPSVPLAPGSAOATNSWTLGCLVKGYPPEVTVTWNSGSLSSGVHFFPAVLQSD 180
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 118 AKTTPSVPLAPGCDTTGSSVTGCLVKGYPPEVTVTWNSGSLSSGVHFFPAVLQSD 177
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 LYTSSSVTPSPSTWSEVTCNVAPASSTKVDKKIVPRDSGSGPSEK 228
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 178 LYTSSSVTVTSTWPSQSITCNVAHPASSTKVDKKIEPR---GPTIK 222
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
S68212
Ig kappa chain (Mab03-1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 29-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 11-Jan-2000
C:Accession: S68212
R:Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka,
FEBS Lett. 375, 273-276, 1995
A:Title: Thermostable peroxidase activity with a recombinant antibody L chain-porphyr
A:Reference number: S68211; MUID:96085223; PMID:7498516
A:Accession: S68212
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-214 <TAK>
A:Cross-references: EMBL:D29668
C:Superfamily: immunoglobulin V region; immunoglobulin homology
```

```
Query Match 26.3%; Score 926; DB 2; Length 214;
Best Local Similarity 83.1%; Pred. No. 2.8e-41;
Matches 177; Conservative 16; Mismatches 14; Indels 6; Gaps 1;

Qy 460 IVMTQPTPSLLYSAGDRVTITCKASQSVND-----VAVYQKPKQSPKLLISYTSRY 513
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 IVMTQPSLLASVGVQKVTMSKSSQSLNSRNKNYLAWYQKPGQSPKLLVYFASTRE 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 514 AGVPRFSGSGYGTDFTLTISVQAEDAAYFCQDYNSPPTFGGKLEIKRADAAPTIV 573
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 SGVPRDFIGSGSGTDFTLTISTVQAEDLADYFCQHYSTPTFGGKLEIKRADAAPTIV 121
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 574 SIFFPSSEQLTGGASVWCFLLNFPKIDNVKWKIDGSRQNGVLSWTDQDSKDSYSTSM 633
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 SIFFPSSEQLTGGASVWCFLLNFPKIDNVKWKIDGSRQNGVLSWTDQDSKDSYSTSM 181
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 634 SFTLTLDKDEYERHNSYTCEATHKTSTSPVKS 666
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 SFTLTLDKDEYERHNSYTCEATHKTSTSPVKS 214
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 10
C89984
enterotoxin P [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89984
R:Kuroda, M.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89984
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <KUR>
A:Cross-references: GB:BA000018; PID:g13701743; PIDN:BA843036.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: sep
C:Superfamily: enterotoxin B
```

```
Query Match 26.1%; Score 918; DB 2; Length 260;
Best Local Similarity 73.0%; Pred. No. 9.3e-41;
Matches 170; Conservative 21; Mismatches 42; Indels 0; Gaps 0;
```


C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-93/Domain: immunoglobulin homology <IMM>

Query Match 25.2%; Score 886.5; DB 2; Length 217;
Best Local Similarity 79.6%; Pred. No. 3.2e-39;
Matches 172; Conservative 14; Mismatches 25; Indels 5; Gaps 1;

QY 461 VMTQTPTSLVSGADRVITCKASQV-----SNDVAVYQKPGQSPKLLISYTSRVAG 515
Db 1 VMTQSPSLPVLSDGASISCRSSQLVHTNGNTYLHWLQKPGQSPKLLIYKVTFRSG 60
QY 516 VPDFSGSGYGTDFTLTSSVQAEAAVFCQDYNSPPTFGGGTKLEIKRADAAPTYSI 575
Db 61 VPDFSGSGSGTDFTKISRVEADLGVVFCQSQTVPVPTFGSGTKLEIKRADAAPTYSI 120
QY 576 FPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSEKQNGVLSWTDQDSDSTYSMS 635
Db 121 FPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSEKQNGVLSWTDQDSDSTYSMS 180
QY 636 TLTLTKDEYERHNSYTCEATHKTSPIVKSFNRE 671
Db 181 TLTLTKDEYERHNSYTCEATHKTSPIVKSFNRE 216

RESULT 20

S68241
Ig kappa chain V region (Mab13-1) - mouse (fragment)
N:Alternate names: immunoglobulin light chain
C:Species: Mus musculus (house mouse)
C:Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S68241; S68214
R:Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.
submitted to the EMBL Data Library, March 1994
A:Description: Specific peroxidase activity by formation of an antibody L-chain-porphyrin
A:Reference number: S68241
A:Accession: S68241
A:Molecule type: mRNA
A:Residues: 1-218 <TAK>
A:Cross-references: EMBL:D29670; NID:G473962; PIDN:BAA06141.1; PID:G473963
R:Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.
FEBS Lett. 375, 273-276, 1995
A:Title: Thermostable peroxidase activity with a recombinant antibody L chain-porphyrin
A:Reference number: S68211; MUID:96085223; PMID:7498516
A:Accession: S68214
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'NI', 3-212 <TAW>
A:Cross-references: EMBL:D29670
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin

Query Match 25.0%; Score 880; DB 2; Length 218;
Best Local Similarity 78.7%; Pred. No. 6.9e-39;
Matches 170; Conservative 13; Mismatches 29; Indels 4; Gaps 1;

QY 460 IVMTQTPTSLVSGADRVITCKASQVSDNDVAVYQKPGQSPKLLISYTSRVAG 515
Db 2 LVLTQSPASLVLGQRATISCRASVSGVYTHWYQKPGQSPKLLISLATNLESG 61
QY 516 VPDFSGSGYGTDFTLTSSVQAEAAVFCQDYNSPPTFGGGTKLEIKRADAAPTYSI 575
Db 62 VPARFSGSGSGTDFTLNIHPVEEDVATYQCQSRELPLTFGAGTKLEIKRADAAPTYSI 121
QY 576 FPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSEKQNGVLSWTDQDSDSTYSMS 635
Db 122 FPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSEKQNGVLSWTDQDSDSTYSMS 181
QY 636 TLTLTKDEYERHNSYTCEATHKTSPIVKSFNRE 671
Db 182 TLTLTKDEYERHNSYTCEATHKTSPIVKSFNRE 217

RESULT 21

S25058
Ig kappa chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jan-2000
C:Accession: S25058
R:Fischer, R.; Voss, A.; Niersbach, M.; Munziker, W.; Hirsch, H.J.; Kreuzaler, F.
submitted to the EMBL Data Library, July 1992
A:Description: Production of a Tobacco mosaic virus (TMV) inactivating neotop specific
A:Reference number: S25057
A:Accession: S25058
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-235 <FIS>
A:Cross-references: EMBL:X67211; NID:G54828; PIDN:CAA47650.1; PID:G54829
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:38-111/Domain: immunoglobulin homology <IMM>

Query Match 24.9%; Score 878.5; DB 2; Length 235;
Best Local Similarity 76.7%; Pred. No. 9e-39;
Matches 171; Conservative 19; Mismatches 32; Indels 1; Gaps 1;

QY 449 LSISLYLTTSTVMTQTPTSLVSGADRVITCKASQVSDNDVAVYQKPGQSPKLLISY 508
Db 13 ISASVLIISRGQIVLTQSPAIMSASPGKVTMTCSASSVSK-MQWYQKSGTSPKRWIYD 71
QY 509 TSSRYAGVPDRFSGSGYGTDFTLTSSVQAEAAVFCQDYNSPPTFGGGTKLEIKRAD 568
Db 72 TSKLASGVPRFSGSGSGTYSLSLTSSMEADAATYTCQWSSNPLTFGAGTKLEIKRAD 131
QY 569 AAPTYSIFPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSEKQNGVLSWTDQDSDK 628
Db 132 AAPTYSIFPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSEKQNGVLSWTDQDSDK 191
QY 629 STYSMSSTLTLTKDEYERHNSYTCEATHKTSPIVKSFNRE 671
Db 192 STYSMSSTLTLTKDEYERHNSYTCEATHKTSPIVKSFNRE 234

RESULT 22

JL0029
Ig kappa chain precursor (RP93) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 11-Jan-2000
C:Accession: JL0029
J.Chien, N.C.; Pollock, R.R.; Desaymard, C.; Scharff, M.D.
J. Exp. Med. 167, 954-973, 1988
A:Title: Point mutations cause the somatic diversification of IgM and IgG2a antiphosph
A:Reference number: JL0029; MUID:88171315; PMID:3127529
A:Accession: JL0029
A:Molecule type: mRNA
A:Residues: 1-225 <CHI>
A:Experimental source: strain BALB/c, cell line RP93 hybridoma cell
A:Note: the authors translated the codon CGG for residue 106 as Pro, ACC for residue
A:Note: the nucleotide sequence shown is inconsistent with authors' translation becau
ect except for four positions shown above
C:Comment: The protein is an anti-phosphorylcholine antibody.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin
F:7-225/Product: Ig kappa chain #status predicted <ILC>
F:7-106/Domain: V region #status predicted <VAR>
F:107-119/Domain: J region #status predicted <JIR>
F:120-225/Domain: C region #status predicted <COR>

Query Match 24.9%; Score 876.5; DB 2; Length 225;
Best Local Similarity 76.4%; Pred. No. 1.1e-38;
Matches 168; Conservative 21; Mismatches 26; Indels 5; Gaps 1;

QY 457 TTSIVMTQTPTSLVSGADRVITCKASQVSDNDVAVYQKPGQSPKLLISYTS 511
Db 5 SSDVLMTQIPLSLPVLSDGASISCRSSQNVHSTGNTYLEWYQKPGQSPNLLIYKISN 64
QY 512 RYAGVPDRFSGSGYGTDFTLTSSVQAEAAVFCQDYNSPPTFGGGTKLEIKRADAAP 571


```
Qy 415 EGSTVSYDLFDAQQGYDPTLLRIYRDNTTISTSTLSLSLYLTYTTSIVMTQTPTSLLSVAG 474
Db 299 -----PD-----VQISWFNVNVEVLTAQTQT----- 319
Qy 475 DRVTITTKASQSVSNDVAWYQKPGSKLLISYSSRYAGVPDRFSGSGYGTDFTLTIS 534
Db 320 -----HREDYNSTR 329
Qy 535 SVQA-----EDAAVYFCQDYNSPTFGGKTKLEIKRADAAPTIVSIFPPSSQQLTSG 586
Db 330 VVSALPIQHDMWSKEFKCKVNNKDLPAPIERTISKIGIVRAPQVILSPPPQLSRK 389
Qy 587 GASVYCFLLNYPKDINVKWKIDGSEQN-----GVLNSWTDODSKDSTYSMSSTLTITK 641
Db 390 DVSLTCLAVGSPEDISVEMTSNGHTEENYKDTAPVLD-----DGSYFIYSKLNMT 442
Qy 642 DEYERHNSYTCEATHK 657
Db 443 SKWKTDSFSCNVRHE 458

RESULT 26
PC4436
monoclonal antibody 13-1 heavy chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jan-2000
C:Accession: PC4436
R:Akashi, S.; Kato, K.; Torizawa, T.; Dohmae, N.; Yamauchi, H.; Kanachi, M.; Harada, A.
Biochem. Biophys. Res. Commun. 240, 566-572, 1997
A:Title: Structural characterization of mouse monoclonal antibody 13-1 against a porphy-
F:22/Disulfide bonds: interchain (to 98) #status predicted
A:Reference number: JC5810; MUID:98063277; PMID:9398605
A:Accession: PC4436
A:Molecule type: protein
A:Residues: 1-444 <AKA>
C:Comment: This catalytic antibody has peroxidase oxidase activity. It is directed again
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:251-320/Domain: immunoglobulin homology <IMM>
F:22/Disulfide bonds: interchain (to 98) #status predicted
F:99/Disulfide bonds: interchain (to 109) #status predicted

Query Match 24.7%; Score 870; DB 2; Length 444;
Best Local Similarity 32.0%; Pred. No. 5.5e-38;
Matches 214; Conservative 71; Mismatches 131; Indels 252; Gaps 15;
Qy 1 EYLOQSGPDLVKPGASVKISKAGSYGFTGYMHVWKVQSGKGLWIG--RINPNNGVT 58
Db 1 EVQXVETGGGLVPGNSLKLCLTSGFTFSNYRHWLQPPGKRLWATVIVKSDNGA 60
Qy 59 LYNQKFKDKATLVDSKSTAYMELRSLTSEDSAYVYCARSTMITNYVNDYWGQGTSTV 118
Db 61 KYAESVRGFTISRDSDSKSSVYLQMNRLREEDTATYYCCRTPWV--YAMDCWGQGTSTV 118
Qy 119 SSKATTPPSVPLAPGSAQAQNSMTLGLCKLVKGYFPEPVTVWNSGSLSSGVHTFPV 178
Db 119 SSKATTPPSVPLAPGSAQAQNSMTLGLCKLVKGYFPEPVTVWNSGSLSSGVHTFPV 178
Qy 179 SLYTLSSSVTVPSWTPSETVTCNVAHPASSTKYDKKIVPRDSDGPGSEKSEINEKDLR 238
Db 179 SLYTLSSSVTVPSWTPSETVTCNVAHPASSTKYDKKIVPRDCG----- 223
Qy 239 KKSELOGTALGNLKOIYYNNSKAITSSEKSDQFNTLLFKGFTGHPWYNDLLVDLGS 298
Db 224 --CKPICITVPEVSSVFIFPPK-----PKDVLITL----- 252
Qy 299 TAATSEYEGSSVDLGAAYGYQCAGGTGPNKTCACMGVGTTLHDNNRLTBKVKPINLWDG 358
Db 253 -----TP-KVFCVVVDIS-----KDDPEVQFSWFVD- 277
Qy 359 KOTTVPIDKVKTSKREV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVGGLI 409
Db 278 ---DVEVHTAQTPREEQFNSTFRSVSELPIMHODWLNKKEFKCRVNSAAPPAPTEK--- 331
Qy 410 VPHSSEGSTVSYVDLPDAQOQYDPTLLRIYRDNTTISTSTLSLSLYLTYTTSIVMTQTPTSL 469
```

```
Db 332 -----TISKT----- 336
Qy 470 LVSAGDRVTITTKASQSVSNDVAWYQKPGSKLLISYSSRYAGVPDRFSGSGYGTDF 529
Db 337 ----- 336
Qy 530 TLTSSVQAEDAAYVFCQDYNSPTFGGKTKLEIKRADAAPTIVSIFPPSSQQLTSGGAS 589
Db 337 -----KGRPKAPOVVTIIPPKEQMAKDVS 361
Qy 550 VVCFLLNYPKDINVKWKIDGSEQN-----GVLNSWTDODSKDSTYSMSSTLTITKDEYERHNS 649
Db 362 LTCMITDFPEDITVEWQWNGOPAE-N-YKNTOPIMDT-DGSYFYYSKLVNOKSNWEAGT 419
Qy 650 YTCEATHK 657
Db 420 FTCSVLHE 427

RESULT 27
A56169
Ig kappa chain V region (clone 23.2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 11-Jan-2000
C:Accession: A56169
R:Monfardini, C.; Kieber-Emmons, T.; Vontfeldt, J.M.; O'Malley, B.; Rosenbaum, H.; God
J. Biol. Chem. 270, 6628-6638, 1995
A:Title: Recombinant antibodies in bioactive peptide design.
A:Reference number: A56169; MUID:95204454; PMID:7896802
A:Accession: A56169
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-210 <MON>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 24.3%; Score 857; DB 2; Length 210;
Best Local Similarity 79.4%; Pred. No. 1e-37;
Matches 166; Conservative 13; Mismatches 26; Indels 4; Gaps 1;
Qy 450 IVMTQPTSLVSGADRVITTCASQSVSND-----VAVYQKPGSKLLISYSSRYAG 515
Db 2 IVLTQSPATLSVSLQRATISCRASKVSSSGSYMHVYQKPGQPKVLIYLASNLESG 61
Qy 516 VPDRESGSGYGTDFTLTSSVQAEDAAYVFCQDYNSPTFGGKTKLEIKRADAAPTIVSI 575
Db 62 VPPRESGSGSGTDFTLNHPVEEEDAATYYCOHSRELPMWTFGGGTRLEIKRADAAPTIVSI 121
Qy 576 FPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSEQNGLNSWTDQDSKDYSSMS 635
Db 122 FPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSEQNGLNSWTDQDSKDYSSMS 181
Qy 636 TLTITKDEYERHNSYTCEATHKSTSPIV 664
Db 182 TLTITKDEYERHNSYTCEATHKSTSPIV 210

RESULT 28
S06084
Ig kappa chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jan-2000
C:Accession: S06084
R:Crowe, J.S.; Smith, M.A.; Cooper, H.J.
Nucleic Acids Res. 17, 7992, 1989
A:Title: Nucleotide sequence of Y3-Ag 1.2.3. rat myeloma immunoglobulin kappa chain c
A:Reference number: S06084; MUID:90016888; PMID:2508067
A:Accession: S06084
A:Molecule type: mRNA
A:Residues: 1-240 <CRO>
A:Cross-references: EMBL:X16129; NID:956457; PIDN:CAA34256.1; PID:956458
C:Superfamily: immunoglobulin V region; immunoglobulin homology
```



```
RESULT 32
JE0244
Ig kappa chain NIG2 precursor - human
C:Species: Homo sapiens (man)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: JE0244
R:Alim, M.A.; Hara, Y.; Hossain, M.S.; Takeda, K.; Yamagata, F.; Yamaki, S.; Kazi, H.; et al.
submitted to JIPID, November 1998
A:Description: A new subgroup of k type light chains (Vkv) identified in cases of AL amyloidosis
A:Reference number: JE0243
A:Accession: JE0244
A:Molecule type: protein
A:Residues: 1-215 <ALI>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 19.9%; Score 702.5; DB 2; Length 215;
Best Local Similarity 62.1%; Pred. No. 1e-29;
Matches 133; Conservative 36; Mismatches 42; Indels 3; Gaps 2;

QY 460 IVWTTPTSLVSGADRVITTCASQSVNDVAVYQKQPGQSPKLLISYTSRYAGVDPDR 519
Db 2 VLVTSPTSLVSPGERATLSCRASQSVHNLAWYQKQPGQAPRLLIYRSTRATGIPAR 61
QY 520 FSGSGYGTDFLTITSSVQAEAAVYFCQDYNS--PPTFGGQTKLEIKRADAAPTVSIFP 577
Db 62 FSGSGGTDFLTITSSLOSEDFALYVCOQ-YNTWPLTFGGGQTKVEIKRTVAAPSVIFP 120
QY 578 PSEQLTSGASVCFVFLNNFYPKDINVKWIDGSRQNGVLNSWTDQDSKDYTSMSSTL 637
Db 121 PSDEQLKSGTASVCLLNNFYPREAKVQKVDNALQSGNSQESVTEQDSKDYTSLSSTL 180
QY 638 TLTKDEYERHNSYTCETHKSTSPIVKSFNRNE 671
Db 181 TLTKADYEAHKYACEVTHQGLSSPVTKSFNRGE 214

RESULT 33
S04845
Ig heavy chain precursor - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 26-Aug-1999
C:Accession: S04845; S05695
R:Ameniya, C.T.; Haire, R.N.; Litman, G.W.
Nucleic Acids Res. 17, 5388, 1989
A:Title: Nucleotide sequence of a cDNA encoding a third distinct xenopus immunoglobulin
A:Reference number: S04845; MUID:89345103; PMID:2503814
A:Accession: S04845
A:Molecule type: mRNA
A:Residues: 1-549 <AME>
A:Cross-references: EMBL:X15114
R:Litman, G.W.
submitted to the EMBL Data Library, April 1989
A:Reference number: S05695
A:Accession: S05695
A:Molecule type: mRNA
A:Residues: 'LC', 3-308, 'H', 310-549, <LIT>
A:Cross-references: EMBL:X15114; NID:g64799; PID:g763031
C:Superfamily: immunoglobulin C region; immunoglobulin homology
A:Keywords: glycoprotein; heterotetramer; immunoglobulin
F:26-109/Domain: immunoglobulin homology <IMM>
F:281.294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.9%; Score 700.5; DB 2; Length 549;
Best Local Similarity 28.0%; Pred. No. 4.2e-29;
Matches 194; Conservative 107; Mismatches 214; Indels 179; Gaps 23;

QY 1 EVLOQSGDPDLVKPGASVKISCKASGYPTGYMHVWVKVQSPGKLEIWRINPNNGVTLY 60
Db 12 DIELVQPSSEIKSPESIKLSCKTSCYFTNTVNIHWIQVQPKGLQWIGRIYPGDADTDY 71
QY 61 NOKFKDKATLVTDKSTTAYMELRSLTSDSAVYYCARSTMTINYMVDYWGQGTSTVSS 120
```

```
Db 72 SSSYQGRCHISTDNPQSTTFLLQNLNKVEDTAIYYCAREGV--GVYFDYWGQGTVMVTVS 129
QY 121 AKTTPSPVYPLAP--GSAQTNSMVTLCCLVKGYPEPVTVTWNSGLSSGVHTTTPAVL- 177
Db 130 AFLHAPSVFPLRPPCGSSS--SDSHVTIGCLTGFUPAPVDVKNVSGSTISGLKNFPVLIQ 188
QY 178 QSDLTSLSSVTPSPSTWPS-ETVTCNVAHPASSTKVDKKIVPRDRSGGSGSESEINEKD 236
Db 189 QSGLFASSQLTIPISDKWAKKSFECNVEHKPTSKVTKQIECQDEPEPIETVEI---- 244
QY 237 LRKKELOQTALGNLKIYYIYNSKAITSEKSADQFLTNTLLFKGFTTGHWPYNOLLVDL 296
Db 245 -----LQG----- 247
QY 297 GSTAATSEYEGSSVDLYGAYGYCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWI 356
Db 248 -----PCASSKSVELLCLITGYA-----PSEIKVHWLL 275
QY 357 DGKQTTVPIDKVKTSKKE-----VTVOELDLQARHYLHGKFGLYNS--DSFGGKVQR 406
Db 276 NQOVNTI SPNSKPKCKEENGTFSSRSKVSVPKED-----WNSEDSYCKVYTH 322
QY 407 GLIVFHS--SEGTVSYDLFDAQGYPTDLLRIYRDNNTTISSTLSISLYLYTTSI--VMTQ 464
Db 323 P--ASHTKTEASTKKC-----DETAI-----TPKVDVLP 350
QY 465 TPTSLVLSAGDVRTITCRASQSVNDVAVYQKQPGQSPKLLISYTSRYAGVDPDRFSGSG 524
Db 351 SPKDLVLT--KEAKYVCVLSRMASTDDLTQVQWSRDKKAL-----AFDSAPEKAYDGT 402
QY 525 YCTDFTLTISVQAEAAVYFCQ---QDYNSPPTFGGQTKLEIKRAD---AAPTVSIFPP 578
Db 403 FTVKSTLISPGDWENKKQFNCKVVPDLPS-----IEKSIQKSDPCTEPTITLLPP 456
QY 579 SSEQLTSGASVCFVFLNNFYPKDINVKWIDGSRQNGVLNSWTD--QDSKDYTSMSSTL 637
Db 457 SDEELRNDPISLICMLKNFRPQDIYVFNKDKGVTLEEDYVMTTTPVLEEEEGFISFSL 516
QY 638 TLTKDEYERHNSYTCETHKSTSPIVKSFNRNE 671
Db 517 TIARSDMWRGATVSCYIAAHNTISQRDIRK--NRGK 549

RESULT 34
JE0242
Ig kappa chain NIG26 precursor - human
C:Species: Homo sapiens (man)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: JE0242
R:Alim, M.A.; Yamaki, S.; Hossain, M.S.; Takeda, K.; Kojima, M.; Takashi, I.; Shinoda
submitted to JIPID, November 1998
A:Description: Structure relationship of kappa type light chains with AL amyloidosis:
A:Reference number: JE0241
A:Accession: JE0242
A:Molecule type: protein
A:Residues: 1-215 <ALI>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:16-91/Domain: immunoglobulin homology <IMM>

Query Match 19.7%; Score 695.5; DB 2; Length 215;
Best Local Similarity 63.1%; Pred. No. 2.3e-29;
Matches 135; Conservative 32; Mismatches 44; Indels 3; Gaps 3;

QY 460 IVWTTPTSLVSGADRVITTCASQSVND--VAVYQKQPGQSPKLLISYTSRYAGVDP 518
Db 2 IVLTQSPGTLSPGERATLSCRASQSVNNYLAHYQKQPGQAPSLIIYDASSRATGIPD 61
QY 519 RFGSGYGTDFLTITSSVQAEAAVYFCQDYNSPP--TPGGGQTKLEIKRADAAPTVSIFP 577
Db 62 RFGSGSGTDFLTITSGLEPDAFYVCOQ--YDRPWTFGGQTKVEIKRTVAAPSVIFP 120
QY 578 PSEQLTSGASVCFVFLNNFYPKDINVKWIDGSRQNGVLNSWTDQDSKDYTSMSSTL 637
```

Db 121 PDEQLKSTASVCLNNFYPREAKVQKVDNALQSGNSQESVTEQDSKSTYLSSTL 180
Qy 638 TLTKEDEYERHNSYTCEATHKTSTSPIVKSFNRNE 671
Db 181 TLSKADYEKHKYACEVTHQGLSSPVTKSFNRGE 214
RESULT 35
B31790
Ig heavy chain V region (17/9) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 23-May-1997
C:Accession: B31790
R:Schulze-Gahmen, U.; Rini, J.M.; Arevalo, J.; Stura, E.A.; Kenten, J.H.; Wilson, I.A.
J. Biol. Chem. 263, 17100-17105, 1988
A:Title: Preliminary crystallographic data, primary sequence, and binding data for an an
A:Reference number: A92686; MUID:89034213; PMID:3182835
A:Accession: B31790
A:Molecule type: mRNA
A:Residues: 1-254 <SCH>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
Query Match 19.4%; Score 683.5; DB 2; Length 254;
Best Local Similarity 63.8%; Pred. No. 1.2e-28;
Matches 146; Conservative 22; Mismatches 48; Indels 13; Gaps 8;
Qy 1 EVQLQSGDPLVKPASKVSKASGYSTGYMHVWVQSPCKGLEWIGRINPNNGVTLY 60
Db 1 EVQLVSGDPLVKPGSKLSCAASGFSSYGMVSRQTPDKRLEWATISNGGGYTY 60
Qy 61 NQKFKDKATLTVDKSTAYMELRSLTSDSAVYVCARSTMITNVMYVMDVWGOGTSTVSS 120
Db 61 PSVVKGRFTISRDNKNTLYLQMSLSKSDSAMYCARRERYDENGFAVWGOGTLVTVSA 120
Qy 121 AKTTPSVVYPLAP--GSAAQTNMWTGLCLVKGYFPEPVTV-TW---NSGSLSSG-VHT 172
Db 121 AKTTPSVVYPLAPVCGXXDTTGSSTGLCLVKGYFPEPVTLTWXXXNSGSLSSGXXVHT 180
Qy 173 FPAVLQS--DLVTLSSSVTVPS--TWP-SETVT-CNVAHPASSTKVDKK 216
Db 181 FPAVLQSXXDLVTLSSSVTVTSSVTPXQSSTFXCNVAHPASSTKVDKK 229
RESULT 36
JE0243
Ig kappa chain NIG93 precursor - human
C:Species: Homo sapiens (man)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: JE0243
R:Allm, M.A.; Hara, Y.; Hossain, M.S.; Takeda, K.; Yamagata, F.; Yamaki, S.; Kazi, H.; T
submitted to JIPID, November 1998
A:Description: A new subgroup of k type light chains (VkV) identified in cases of AL amy
A:Reference number: JE0243
A:Accession: JE0243
A:Molecule type: protein
A:Residues: 1-215 <ALI>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:16-90/Domain: immunoglobulin homology <IMM>
Query Match 19.3%; Score 679.5; DB 2; Length 215;
Best Local Similarity 61.0%; Pred. No. 1.6e-28;
Matches 130; Conservative 33; Mismatches 49; Indels 1; Gaps 1;
Qy 460 IWMTQPTSLVYAGDRVTITCKASQSVNDVAVYQKPGQPKLLISYTSRYAGVPDR 519
Db 2 IWMTQSPATLSVSPGERATLSCRASQSVATNVVMYMQLGQAPRELLIYDASTRATGVPAR 61
Qy 520 FSGSGYGTDTLTITSSVQAEDAAYVFCQDYNSTPTFGGTTKLEIKRADAPTYSIFPP 578
Db 62 FSGSGGTFTLTITSSLOSEFAIYYCQHNNAWPPPTFGGTTKRTVAAPSVFIAPP 121

Qy 579 SSEQLTSGASVYVCFLLNFPKIDINVKWKIDGERONGVLNSWTDQDSKSDSTYSKSTLT 638
Db 122 SDEQLKSGTASVVCLLNNFYPREAKVQKVDNALQSGNSQESVTEQDSKSDSTYLSSTLT 181
Qy 639 LTKDEYERHNSYTCEATHKTSTSPIVKSFNRNE 671
Db 182 LSKADYEKHKYACEVTHQGLSSPVTKSFNRGE 214
RESULT 37
S33161
Ig kappa chain - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000
C:Accession: S33161
R:Foley, R.C.; Beh, K.J.
submitted to the EMBL Data Library, July 1990
A:Description: Isolation and characterisation of sheep kappa light chain cDNA.
A:Reference number: S33161
A:Accession: S33161
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-230 <FOL>
A:Cross-references: EMBL:X54110; NID:g297103; PIDN:CAA38046.1; PID:g1364221
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:143-212/Domain: immunoglobulin homology <IMM>
Query Match 19.2%; Score 675; DB 2; Length 230;
Best Local Similarity 59.4%; Pred. No. 2.9e-28;
Matches 126; Conservative 37; Mismatches 49; Indels 0; Gaps 0;
Qy 460 IWMTQPTSLVYAGDRVTITCKASQSVNDVAVYQKPGQPKLLISYTSRYAGVPDR 519
Db 18 IQVTPSPSLASLTERVSTICRTSQSVSNYLNWYQKPGQAPKLLIYATRLHTDVPDR 77
Qy 520 FSGSGYGTDTLTITSSVQAEDAAYVFCQDYNSTPTFGGTTKLEIKRADAPTYSIFPP 579
Db 78 FSGSGGTDTLTITSNLEANDATYYCLQYESTPLAFGGGTNVEIKRSDAQSPVLFKPS 137
Qy 580 SQQLTSGASVYVCFLLNFPKIDINVKWKIDGERONGVLNSWTDQDSKSDSTYSMSSTLT 639
Db 138 EQQLTGTVTSVVCLVNDVFKDINVKVGVTONSFQNSFTDQDSKSTYLSSTLT 197
Qy 640 TKDEYERHNSYTCEATHKTSTSPIVKSFNRNE 671
Db 198 SSSEIQSHNAYACEVSHKSLPTALVKSPNKNE 229
RESULT 38
S22080
Ig heavy chain precursor (B/MT.4A.17.H5.A5) - bovine
N:Alternate names: Ig gamma-1 chain C region (clone 8.10)
C:Species: Bos primigenius taurus (cattle)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S22080; S06610; A31303
R:Sanders, P.G.
submitted to the EMBL Data Library, November 1991
A:Reference number: S22080
A:Accession: S22080
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-470 <SAN>
A:Cross-references: EMBL:X62916; NID:g439; PIDN:CAA44699.1; PID:g440
R:Symons, D.B.A.; Clarkson, C.A.; Beale, D.
Mol. Immunol. 26, 841-850, 1989
A:Title: Structure of bovine immunoglobulin constant region heavy chain gamma 1 and g
A:Reference number: S06610; MUID:90097956; PMID:2513487
A:Accession: S06610
A:Molecule type: DNA
A:Residues: 142-470 <SYM>
A:Cross-references: EMBL:X16701

C>Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 16-Jul-1999
C:Accession: A49444
R:Saul, F.A.; Poljak, R.J.
Proteins 14, 363-371, 1992
A:Title: Crystal structure of human immunoglobulin fragment Fab New refined at 2.0 A resolution
A:Reference number: A49444; MUID:93066153; PMID:1438175
A:Accession: A49444
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-220 <SAU>
A>Note: sequence modified after extraction from NCBI backbone
A>Note: this sequence report includes corrections based on crystal structure refinement
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:137-202/Domain: immunoglobulin homology <IMM>

Query Match 17.3%; Score 611; DB 2; Length 220;
Best Local Similarity 54.3%; Pred. No. 5.6e-25;
Matches 120; Conservative 36; Mismatches 61; Indels 4; Gaps 3;

QY 1 EVLOQSGPDLVKPGASVKISKASCYSFTGYMHVWVKSPGKGLGWIGRINPNNGVTLY 60
DB 1 QVLESGGGLVLRPSQTLSTCVSTSFDDIYTWVRQPGKLEWIGYVF-YTGTTLL 59
QY 61 NQKFKDKATLTVDKSTTAYMELRLSTSDSAVYICARSTMITNYMYDYGQTSVTSS 120
DB 60 DPSLRGRVTMLVNTSKNQSRLSSVTAADTAVYICARNLIAGG--IDVWGQSLVTSS 117
QY 121 AKTPPSVPLAPGSAQNTNSWYTLCLVKGYFPEPVTYTWNSGSLSSGWHFPAVLQSD 180
DB 118 ASKPGSVPLAPSSKSTSGGTAAALGCKYKDFPEPVTYTWNSGALTSVGTFFPAVLQSS 177
QY 181 -LYLTSSVTPSPSTPSETVTCNVAHPASSTKVDKIKVPR 220
DB 178 GLYSLSSVTPSPSSLTGTQYICNVNHHKPSNTKYDKRKEPK 218

RESULT 45
A33953
enterotoxin D precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 15-Oct-1999
R:Bayles, K.W.; Iandolo, J.J.
J. Bacteriol. 171, 4799-4806, 1989
A:Title: Genetic and molecular analyses of the gene encoding staphylococcal enterotoxin
A:Reference number: A33953; MUID:89359112; PMID:2549000
A:Accession: A33953
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-258 <BAV>
A:Cross-references: GB:M28521; NID:g1492109; PIDN:AAB06195.1; PID:g758691
C:Superfamily: enterotoxin B

Query Match 17.3%; Score 611; DB 2; Length 258;
Best Local Similarity 51.1%; Pred. No. 6.9e-25;
Matches 118; Conservative 35; Mismatches 78; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLKKSELGATLGNLQKIIYYNSKAITSEKSADQFLNTLLFKGFTFG 285
DB 26 NEMIDSVKEKLHKKSELSSTALNNKHSYADKNPIIGENKSTGDFLENTLLYKFFTD 85
QY 286 HPWYNLLVDLSTAATSEYEGSSVDLYGAYGYQCAGTPNKTCMYGGVTLHDNRILT 345
DB 86 LINFEDLLINFNSEKMAQHFKSKNVDVYPIRVSINCYGGEIDRTACTYGGVTPHEGNLKL 145
QY 346 EEKKVPINLWDGKQTPVPIDKVKTSKKEVTVQELQARHLHGKFGLYNSDSFGKQVQ 405
DB 146 ERKKIPINLWINGVQREVSIDKQVTDKKNVTYQELDAQARRYLQDKLYLNNDTLGGKIQ 205
QY 406 RGLVPHSSGEGSTVSDLEDAQGPDPDILLRIYRDNTTISSTLSLSLYL 456
DB 206 RGKIFEDSDGSKVSDLEFDVKGDFPEKQIRIYSNKTLSLTHELHIDYLY 256

RESULT 46
S14683
Ig mu chain precursor, membrane-bound (clone 201) - human
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Jul-1999
C:Accession: S14683; S08047
R:Friedlander, R.M.; Nussenzweig, M.C.; Leder, P.
Nucleic Acids, Res. 18, 4278, 1990
A:Title: Complete nucleotide sequence of the membrane form of the human IgM heavy chain
A:Reference number: S14683; MUID:90332450; PMID:2115996
A:Accession: S14683
A:Molecule type: mRNA
A:Residues: 1-627 <FRI>
A:Cross-references: EMBL:X17115; NID:g33450; PIDN:CAA34971.1; PID:g33451
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin; membrane protein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-627/Product: Ig mu chain #status predicted <MAT>
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 16.9%; Score 595.5; DB 2; Length 627;
Best Local Similarity 26.8%; Pred. No. 1.3e-23;
Matches 190; Conservative 100; Mismatches 209; Indels 211; Gaps 27;

QY 1 EVLOQSGPDLVKPGASVKISKASCYSFTGYMHVWVKSPGKGLGWIGRINPNNGVTLY 60
DB 20 QVOLVOSGAEVKRPSSSVKSCAGTGFSSVAISWVRQAPGOGLEWGGIIPFGTANY 79
QY 61 NQKFKDKATLTVDKSTTAYMELRLSTSDSAVYICARSTMITNYV-----M 107
DB 80 AQKFGQRTVITADESTSTAYMELSLRSEDYAVYICAKTGILGYPSSGWHFPAVLQSD 139
QY 108 DYWGQTSVTVSSAKTTPPSVYPL--APGSAQNTNSWYTLCLVKGYFPEPVTYTW--N 162
DB 140 DVWGQTTTVTVSSGASAPTLFPLVSCNSPDTSS--VAVGCLAQDFLPDSITFSWKYN 198
QY 163 SGSLSSGVHTFPAVLQSDLYTLSSSVTPSS---TWPSETVTCNVAHPASS----- 210
DB 199 NSDISS--TRGFPVSLRGKGKIAATQVLLPSKDVMOGTDEHVYCKVOHPNGNKEKNVPLPV 257
QY 211 -----TKVDKIKVPRDS--GGPSEKSENEKDLRKKSELGATLGNLQKIIYYNSKAIT 263
DB 258 IAEPLPKVSFVPPRDPFGFNPRSKSLICQATGSPRQIOWSLREGQV----- 308
QY 264 SSEKSADQFLNTLLFKGFTGHPWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQCAG 323
DB 309 -----GSGVTTDQVQAEAKE-----SG 325
QY 324 GTPNKTCMYGGVTLHDNRILTEEKVPIINLWDGKQTPVPIDKVKTSKKEVTVQELDLO 383
DB 326 PTYKVT---STLTIKESD-----WLSOSMFTCRVD----- 353
QY 384 ARHYLHGKFGLYNSDSFGKQVQRLIVFHSSEGSTVSYDLFDAQGOYPDILLRIYRDNTT 443
DB 354 -----HRGL-TFOQNASSMCVPD-----QDTARV-----A 379
QY 444 ISSTLSLSLYLYTTSIVMTQPTSLLSAGDRVTITCKASQSVNSDVAVYQOKPGQSPK 503
DB 380 IPPSPASIFL---TKSTKLTLCLVTOL--TTYDSVTI-----SWTRQN-GEAVK 421
QY 504 LLISYTSRRYAGVPDRFSGSGYGTDTLLTSSVQAEADAAYVFCQODYNSPPTF----- 556
DB 422 ---THTNISESHPNATFSAVG-----EASI---CEDDWSNGERTCTVTHT 461
QY 557 --GGGTKLKIKRADAA---PTVSIFPPSPSEQLT--SGGASVVCFLNNPKDINVKWKID 609
DB 462 DLPSPLKQIISRPGKVALHRPDYVLLPAREQLNLRRESATITCLYTGFSFADVFQVMQR 521
QY 610 GS--ERQNGVLNSWTDQDSKSTYSMSSTLTUTKDEYERHNSYTCETHK 657
DB 522 GQPLSPEKYVTSAPWPEQAPGRYFAHSILTVSEEWNTGETYTCVVAHE 571

F:143-212/Domain: immunoglobulin homology <IMM>

```

RESULT 47
PT0219
Ig kappa chain V-C region (PLC18) - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 11-Jan-2000
C:Accession: PT0219
R:Lammers, B.M.; Beaman, K.D.; Kim, Y.B.
Mol. Immunol. 28, 877-880, 1991
A:Title: Sequence analysis of porcine immunoglobulin light chain cDNAs.
A:Reference number: PT0219; MUID:91342694; PMID:1715030
A:Accession: PT0219
A:Molecule type: mRNA
A:Residues: 1-178 <LAM>
A:Cross-references: GB:M59321; NID:q164508; PIDN:AAA03520.1; PID:g164509
A:Experimental source: spleen, strain Minnesota Miniature
A:Note: The authors translated the codon CTC for residue 141 as Ser
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-70/Domain: V region (fragment) <VRG>
F:12-18/Region: complementarity-determining 1
F:19-51/Region: framework 1
F:52-60/Region: complementarity-determining 2
F:61-70/Region: framework 2
F:71-178/Domain: C region <CRG>
F:96-156/Disulfide bonds: #status predicted
F:176/Disulfide bonds: interchain #status predicted

Query Match 16.8%; Score 591; DB 2; Length 178;
Best Local Similarity 64.6%; Pred. No. 4.7e-24;
Matches 113; Conservative 23; Mismatches 39; Indels 0; Gaps 0;

QY 497 KPGQSPKLLISYTSRYAGVPDRSGSGVGTDTLTITSSVQAEADAAVFCQDYNSPPNF 556
Db 1 KPGQSPQLLIVASDRASGVPDRFGSGSGTDTLTIKINSVEADAGVYCHQKFEPRTF 60

QY 557 GGGTKLEIKRADAAPTVSIFPPSPSEQLTSGGASVVCFLNFPKIDINVKWKIDGSGRQNG 616
Db 1 GGGTKLEIKRADAKSVIFPPSPSEQLTSGGASVVCFLNFPKIDINVKWKIDGSGRQNG 616

QY 617 VLNSWTDQSKDSTYSMSSTLTLTDEYERHNSYTCETHKSTSPIVKSFNRNE 671
Db 121 HPDSVTEQDSKSTYSLSLSTLPTSQYLSHLNLSYCEVTHKTLASPLVTSFNRNE 175

RESULT 48
A20969
Ig kappa chain precursor V-J-C regions - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 21-Jan-2000
C:Accession: A20969; A25448
R:McCartney-Francis, N.; Skurla Jr., R.M.; Mage, R.G.; Bernstein, K.B.
Proc. Natl. Acad. Sci. U.S.A. 81, 1794-1798, 1984
A:Title: Kappa-chain allotypes and isotypes in the rabbit: cDNA sequences of clones encoding expression.
A:Reference number: A20969; MUID:84170388; PMID:6424124
A:Accession: A20969
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-229 <MCC>
A:Cross-references: GB:K01359; NID:q165373; PIDN:AAA31334.1; PID:g165374
R:Akimenko, M.A.; Marilame, B.; Rougeon, F.
Proc. Natl. Acad. Sci. U.S.A. 83, 5180-5183, 1986
A:Title: Evolution of the immunoglobulin kappa light chain locus in the rabbit: evidence
A:Reference number: A94110; MUID:86259753; PMID:3088570
C:Contents: Ig kappa-1 chain, 69 allotype, J-K1.2 segment
A:Accession: A25448
A:Molecule type: DNA
A:Residues: 111-123 <AKI>
A:Cross-references: GB:M14067; GB:M14062; GB:M14063; GB:M14064; GB:M14065; GB:M14066; NT
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
```

```

Query Match 16.2%; Score 570.5; DB 2; Length 229;
Best Local Similarity 53.0%; Pred. No. 7.4e-23;
Matches 114; Conservative 35; Mismatches 63; Indels 3; Gaps 3;

QY 459 SIVMTQTPTSLVSAGDRVTITCKASQSVSNDAVYQQRPGQSPKLLISYTSRVAAGVPD 518
Db 15 ALVMTQTPASVSAVGGTVTIKQASENIYSSLAWYQQRPGPKLLIYGASTLASGVPS 74

QY 519 RFSGSYGTDFTLTITSSVQAEADAAVFC-QQDYNSPPTFGGKLEIK-RADAAPTSTP 576
Db 75 RFKGRSGTEYTLTISGVQREDAATYYICLGSDDSTAFGGTEILEILCDPPIAPTPLVF 134

QY 577 PPSSEQLTSGGASVVCFLNFPKIDINVKWKIDGSGRQNGVLNSWTDQSKDSTYSMSST 636
Db 135 PPSAQLTETVTIIVCVANKFRPNDITVTWKVDDLEIQQSGIENSTTPQSPEDCTYNLSST 194

QY 637 LTLTKDEYERHNSYTCETHKSTSPIVKSFNRNE 671
Db 195 LSLTRAQYNSHSYTCVVH-NSGSAIVQSFNRGD 228

RESULT 49
S29594
Ig gamma chain (WM65) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S29594
R:Seymour, R.
submitted to the EMBL Data Library, February 1991
A:Reference number: S29593
A:Accession: S29594
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-178 <SEY>
A:Cross-references: EMBL:X57857; NID:g52590; PIDN:CAA40992.1; PID:g52591
C:Keywords: immunoglobulin

Query Match 16.0%; Score 562; DB 2; Length 178;
Best Local Similarity 73.5%; Pred. No. 1.5e-25;
Matches 108; Conservative 13; Mismatches 24; Indels 2; Gaps 1;

QY 1 EVQLQQSGDPLVKPGASVKISCKASGYFTGYMHVWKQSPGKLEWIGRINPNNGVTLY 60
Db 14 EVQLQQSGAEVLKPGASVKLSCTASGFNIKDTYIHHVKQRPKGLEWIGRIDPANGYTEY 73

QY 61 NQKFKDKATLTVDKSTTTAYMELRLSTSDSAVYYCARSTMITNYYMDYWGQGSTVTSS 120
Db 74 DKFKGKATITADTSTNTAYLQLSLSTSEDATVYYCTGNYA--YGM DYWGQGSTVTSS 131

QY 121 AKTTPPSVYPLAPGSAQAQTNMVTLCG 147
Db 132 AKTTPPSVYPLAPGSAQAQTNMVTLCG 158

RESULT 50
I54782
gene Pvt-1a/Ig-Ck protein - mouse (fragment)
C:Species: Mus sp. (mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 02-Sep-2000
C:Accession: I54782
R:Huppi, K.; Siwarski, D.
Int. J. Cancer 59, 848-851, 1994
A:Title: Chimeric transcripts with an open reading frame are generated as a result of
A:Reference number: I54782; MUID:95080867; PMID:7989128
A:Accession: I54782
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-126 <RES>
A:Cross-references: GB:S76258; NID:g913277; PIDN:AAB32752.1; PID:g913278
C:Genetics:
A:Gene: Pvt-1a/Ig-Ck
```

C:Superfamily: pre-B cell omega light chain; immunoglobulin homology

Query Match 15.9%; Score 561; DB 2; Length 126;
Best Local Similarity 93.0%; Pred. No. 1.le-22;
Matches 107; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 557 GGGTKLEIKRADAPTIVSIFPPSSQLTSGGASVVCFLNNFYPKDINVKWKIDGSRQNG 616
Db 11 GAAAKVKSIIWADAAPTIVSIFPPSSQLTSGGASVVCFLNNFYPKDINVKWKIDGSRQNG 70
QY 617 VLNSWTDQDSKDSYMSSTLTLTKDYEYRHSYTCETHKSTSPIVKSFNRNE 671
Db 71 VLNSWTDQDSKDSYMSSTLTLTKDYEYRHSYTCETHKSTSPIVKSFNRNE 125

Search completed: January 8, 2003, 11:58:57
Job time : 32.2464 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:52:22 ; Search time 13.3657 Seconds
(without alignments)
2085.338 Million cell updates/sec

Title: US-09-900-766-1
Perfect score: 3522
Sequence: 1 EVQLQSGPGLVKPGASVKI.....EATHKTSTSPVKSPNRNES 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	1107	31.4	257	1	ETXE_STAAU	P12993 staphylococ
2	948	26.9	257	1	ETXA_STAAW	P13163 staphylococ
3	611	17.3	258	1	ETXD_STAAU	P20723 staphylococ
4	557	15.8	324	1	GCI_MOUSE	P01868 mus musculu
5	557	15.8	393	1	GCIM_MOUSE	P01869 mus musculu
6	556	15.8	106	1	KAC_MOUSE	P01837 mus musculu
7	550	15.6	326	1	GCI_RAT	P20759 rattus norv
8	539	15.3	322	1	GCA_RAT	P20760 rattus norv
9	491.5	14.0	330	1	GCAA_MOUSE	P01863 mus musculu
10	491.5	14.0	399	1	GCAM_MOUSE	P01865 mus musculu
11	481.5	13.7	117	1	HV12_MOUSE	P01756 mus musculu
12	481	13.7	139	1	HV07_MOUSE	P01751 mus musculu
13	480.5	13.6	117	1	HV13_MOUSE	P01757 mus musculu
14	480	13.6	118	1	HV31_MOUSE	P06330 mus musculu
15	479	13.6	336	1	GCB_MOUSE	P01866 mus musculu
16	479	13.6	405	1	GCBM_MOUSE	P01867 mus musculu
17	472	13.4	106	1	KACA_RAT	P01836 rattus norv
18	467	13.3	335	1	GCAB_MOUSE	P01864 mus musculu
19	458.5	13.0	120	1	HV03_MOUSE	P01747 mus musculu
20	457	13.0	106	1	KACB_RAT	P01835 rattus norv
21	454	12.9	137	1	HV11_MOUSE	P01755 mus musculu
22	450.5	12.8	140	1	HV02_MOUSE	P01746 mus musculu
23	446	12.7	333	1	GCB_RAT	P20761 rattus norv
24	437.5	12.4	138	1	HV48_MOUSE	P03980 mus musculu
25	437.5	12.4	329	1	GCC_RAT	P20762 rattus norv
26	433	12.3	120	1	HV30_MOUSE	P06329 mus musculu
27	425	12.1	117	1	HV09_MOUSE	P01753 mus musculu
28	423	12.0	149	1	KV5A_MOUSE	P01633 mus musculu
29	422	12.0	117	1	HV14_MOUSE	P01758 mus musculu
30	420.5	11.9	136	1	HV15_MOUSE	P01759 mus musculu
31	417	11.8	117	1	HV04_MOUSE	P01748 mus musculu
32	415	11.8	117	1	HV10_MOUSE	P01754 mus musculu
33	414	11.8	117	1	HV49_MOUSE	P06328 mus musculu

RESULT 1				
ID	ETXE_STAAU	STANDARD;	PRT;	257 AA.
AC	P12993;			
DT	01-JAN-1990	(Rel. 13, Created)		
DT	01-JAN-1990	(Rel. 13, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Enterotoxin type E precursor (SEE).			
GN	ENTE.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 28-74.			
RC	STRAIN-MJB265;			
RX	MEDLINE=88257005; PubMed=3384800;			
RA	Couch J.L., Soltis M.T., Betley M.J.;			
RT	"Cloning and nucleotide sequence of the type E staphylococcal enterotoxin gene.";			
RT	J. Bacteriol. 170:2954-2960(1988).			
RL	[2]			
RP	3D-STRUCTURE MODELING.			
RX	MEDLINE=96022987; PubMed=7552730;			
RA	Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;			
RT	"Residues defining V beta specificity in staphylococcal enterotoxins.";			
RT	Nat. Struct. Biol. 2:680-686(1995).			
CC	-!- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary for the toxin interaction with MHC class II (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION STAPHYLOCOCCAL FOOD POISONING SYNDROME.			
CC	-!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M21319; AAA26617.1; -.			
DR	PIR; A28179; A28179.			
DR	PDB; 1SEE; 15-OCT-95.			
DR	InterPro; IPR001961; Staph/Strep_toxin.			
DR	Pfam; PF01123; Staph_strep_toxin; 1.			
DR	Pfam; PF02876; Staph_strep_toxin; 1.			
DR	PRINTS; PR00279; BACTRLTOXIN			
DR	PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.			
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.			
KW	Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;			
3D-structure.	1	27		
FT	CHAIN	28	257	ENTEROTOXIN TYPE E.

P01750 mus musculu
P01862 cavia porce
P06313 homo sapien
P06314 homo sapien
P01749 mus musculu
P22436 mus musculu
P03987 mus musculu
P01857 homo sapien
P01745 mus musculu
P01861 homo sapien
P01825 homo sapien
P01634 mus musculu

ALIGNMENTS

```
FT METAL 211 211 ZINC (BY SIMILARITY).
FT METAL 249 249 ZINC (BY SIMILARITY).
FT METAL 251 251 ZINC (BY SIMILARITY).
SQ SEQUENCE 257 AA; 29358 MW; 27EDA94B97770CE3 CRC64;

Query Match 31.4%; Score 1107; DB 1; Length 257;
Best Local Similarity 89.7%; Pred. No. 9.5e-62;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKSELOCTALGNLKOIYYNSKAITSEKSADOFLENTLFLKGFFTG 285
DB 25 SEKSEINEKDLRKSELOCTALGNLKOIYYNEKAITENKESDDQFLENTLFLKGFFTG 84
QY 286 HPWYNLLVDLGSSTAATSEYSGSVLDLYGAYGYOCAGGTPNKATCMYGGVTLHDNNRLT 345
DB 85 HPWYNLLVDLGSKDATNKYKGGKVDLYGAYGYOCAGGTPNKATCMYGGVTLHDNNRLT 144
QY 346 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGKQVQ 405
DB 145 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGKQVQ 204
QY 406 RGLIVHSSEGSVSYDLDAQOGQYPTLLRIYRDNTTISSTSLSLVLYTT 458
DB 205 RGLIVHSSEGSVSYDLDAQOGQYPTLLRIYRDNTTINSENLHIDLXYTT 257

RESULT 2
ETXA_STAAW STANDARD; PRT; 257 AA.
AC FI1363;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type A precursor (SEA).
GN ENTA OR MW1889.
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=120434378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RT Lancet 359:1819-1827(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PRI337;
RX MEDLINE=88086892; PubMed=3335483;
RA Bettley M.J., Mekalanos J.J.;
RT "Nucleotide sequence of the type A staphylococcal enterotoxin gene.";
RT J. Bacteriol. 170:34-41(1988).
RN [3]
RP SEQUENCE OF 25-257.
RX MEDLINE=87222293; PubMed=3584106;
RA Huang I.-Y., Hughes J.L., Bergdoll M.S., Schantz E.J.;
RT "Complete amino acid sequence of staphylococcal enterotoxin A.";
RT J. Biol. Chem. 262:7006-7013(1987).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=95354648; PubMed=7628431;
RA Schad E.M., Zaitseva I., Zaitsev V.N., Dohlsten M., Kalland T.,
RA Schlievert P.M., Ohlendorf D.H., Svensson L.A.;
RT "Crystal structure of the superantigen staphylococcal enterotoxin
RT type A.";
RL EMBO J. 14:3292-3301(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=97113025; PubMed=8943278;
```

```
RA Sundstroem M., Hallen D., Svensson A., Schad E., Dohlsten M.,
RA Abrahmsen L.;
RT "The Co-crystal structure of staphylococcal enterotoxin type A with
RT Zn2+ at 2.7-A resolution. Implications for major histocompatibility
RT complex class II binding.";
RL J. Biol. Chem. 271:32212-32216(1996).
RN [6]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=96022987; PubMed=7552730;
RA Swaminathan S., Furey W.F. Jr., Fletcher J., Sax M.;
RT "Residues defining V beta specificity in staphylococcal
RT enterotoxins.";
RL Nat. Struct. Biol. 2:680-686(1995).
RN [7]
RP COMPARISON OF STRUCTURE OF SEA AND SEC2.
RX MEDLINE=97334373; PubMed=9191070;
RA Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;
RT "A structural and functional comparison of staphylococcal
RT enterotoxins A and C2 reveals remarkable similarity and
RT dissimilarity.";
RL J. Mol. Biol. 269:270-280(1997).
CC -!- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
CC for the toxin interaction with MHC class II.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY A BACTERIOPHAGE.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP004828; BAB95754.1;
DR EMBL: M18970; AAX26681.1;
DR PIR: A28664; A28664.
DR PIR: A29566; A29566.
DR PDB: 1ESF; 11-JUL-96.
DR PDB: 1SEA; 15-OCT-95.
DR PDB: 1SXT; 19-NOV-97.
DR InterPro: IPR001961; Stap/Strep toxin.
DR Pfam: PF01123; Stap_Strp_toxin; 1.
DR Pfam: PF02876; Stap_Strp_tox_C; 1.
DR PRINTS: PR00279; BACTRLTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superaantigen; Metal-binding; Zinc;
KW 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 257 ENTEROTOXIN TYPE A.
FT DISULFID 120 130
FT METAL 211 211 ZINC.
FT METAL 249 249 ZINC.
FT METAL 251 251 ZINC.
FT CONFLICT 242 242 T -> S (IN REF. 3).
SQ SEQUENCE 257 AA; 29669 MW; ADEBF5BCA1F14677 CRC64;

Query Match 26.9%; Score 948; DB 1; Length 257;
Best Local Similarity 76.4%; Pred. No. 6.4e-52;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKSELOCTALGNLKOIYYNSKAITSEKSADOFLENTLFLKGFFTG 285
DB 25 SEKSEINEKDLRKSELOCTALGNLKOIYYNEKAITENKESDDQFLENTLFLKGFFTG 84
QY 286 HPWYNLLVDLGSSTAATSEYSGSVLDLYGAYGYOCAGGTPNKATCMYGGVTLHDNNRLT 345
DB 85 HPWYNLLVDLGSKDATNKYKGGKVDLYGAYGYOCAGGTPNKATCMYGGVTLHDNNRLT 144
```

	Best Local Similarity	51.1%, Pred. No. 4.5e-31;			
	Matches 118; Conservative	35; Mismatches 78; Indels	0; Gaps	0;	
QY	226	SEKSEEINERDLRKSELQGTALCNLKIQIYYVNSKAITSSEKSADQLTNTLLFKGPFITG	285		
		: : : : :	:	:	:
Dd	26	NENIDSVKEHLKKSELSSTALNNMKHSYADKNPIIGENKSGDQFLENTLLVKKFEFD	85		
		: : : : :	:	:	:
QY	286	HPWYNLDLVGLGSTAATSETVEGSSVDLYGAYGYQCAGGTPNKTCAMYGVTTLHDNNRLT	345		
		: : : : :	:	:	:
Dd	86	LINFEDLLIFNFSKEMAHQFKNVDDVPYRISYNCYGGEIDRTACTYGVGPHEGNKLK	145		
		: : : : :	:	:	:
QY	346	EKKVPINLWDGQTVPIDIKVYKTSKEVTVQELDLQAARHYLHGKGLVNSDSFGKGKVQ	405		
		: : : : : :	:	:	:
Dd	146	ERKKIPINLWINGVQEVSIDKVTQDKNVTVELDQAARRYLQKDILKLYNNDTLGGKIQ	205		
		: : : : : :	:	:	:
QY	406	RGLIVFHSSGEGSVSYDLFDAQQGYPTLLRIYRDNTTTISTSLUSILSILYL	456		
		: : : : :	:	:	:
Dd	206	RGKIEFPSSDGSKSVSYDLFDVGDFPEPKQIRIYSDNKTLTSTEHLHIDIYLY	256		
		: : : : :	:	:	:

[illegible]

RESULT 4			
GCI_MOUSE	STANDARD;	PRT;	324 AA.
ID	GCI_MOUSE		
AC	P01868;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	Ig gamma-1 chain C region.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurcoepathi; Muridae; Mus		

OC mammalia; eutheria; rodentia; sciurognathini; muridae; murinae; mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=80045036; Pubmed=115993;
RA Honjo T.; Obata M.; Yamawaki-Kataoka Y.; Kataoka T.; Kawakami T.;
RA Takahashi N.; Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene";
RL Cell 18:559-568(1979).

RN [2] SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
 RP MEDLINE=80202559; PubMed=6769752;
 RX Obata M, Yamawaki-Ikataoka Y, Takahashi N., Kataoka T., Shimizu A.,
 RA Mano Y, Seisawa J, Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
 PA Mano Y, Seisawa J, Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,

Mano I., Setoian J.G., Peeterlin B.M., Leder P., Honjo T.:
"Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
cloned in a bacterial plasmid.";
Gene 9:87-97(1980).
[3]
SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
MEDLINE=80012837; PubMed=113776;
Rogers J., Clarke P., Salsner W.:
"Sequence analysis of cloned cDNA encoding part of an immunoglobulin

RT RT heavy Chain.";
RN RN Nucleic Acids Res. 6:3305-3321(1979).
RN RN [4]
RN RN SEQUENCE (MYELOMA PROTEIN MOPC 21).
RX RX MEDLINE=78242288; PubMed=98524;
RX RX Agetugbo K.;
RT RA "Evolution of immunoglobulin subclasses. Primary structure of a
RT RT murine myeloma gammal chain.";
RT RT J. Biol. Chem. 253:6068-6075(1978).
RN RN [5]
RN RN DISULFIDE BONDS (MOPC 21).
RX RX MEDLINE=7300889; PubMed=5073237;
RX RX Svasti J., Milstein C.;
RA RA "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RN RN Biochem. J. 126:837-850(1972).
RN RN
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).


```

DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00407; Igc1; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 106 106
FT DISULFID 106 106
SQ SEQUENCE 106 AA; 11778 MW; 4B51FF5EF49BAEB5 CRC64;

Query Match 15.8%; Score 556; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.4e-28;
Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 ADAAPTSTVFPPSSQLTSGGASVYVCFNNFYPKDKINVKWKIDGSRQNGVLSNWTDDQS 626
DB 1 ADAAPTSTVFPPSSQLTSGGASVYVCFNNFYPKDKINVKWKIDGSRQNGVLSNWTDDQS 60
QY 627 KDSTYSMSSTLTITKDEYERHNSYTCETHKTSTSPIVKSFNRNE 671
DB 61 KDSTYSMSSTLTITKDEYERHNSYTCETHKTSTSPIVKSFNRNE 105

RESULT 7
GCL_RAT
ID GCL_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-1 chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR HSP; P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00407; Igc1; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 112
FT DOMAIN 113 219
FT DOMAIN 220 326
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 111 111
FT DISULFID 140 200
FT DISULFID 246 306
FT CARBOHYD 176 176
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 15.6%; Score 550; DB 1; Length 326;
Best Local Similarity 27.7%; Pred. No. 3.6e-27;
Matches 150; Conservative 64; Mismatches 90; Indels 238; Gaps 14;

QY 121 AKTTPSVVPLAPGSAQAQNSMVTGLCLVKYGYPPEVPTVWNSGSLSGVHTFPAVLQSD 180
DB 1 AETAPSVVPLAPGALKNSMVTGLCLVKYGYPPEVPTVWNSGSLSGVHTFPAVLQSG 60

```

```

QY 181 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVDKKIVPRDGGPSEKSEINEKDLRKK 240
DB 61 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVDKKIVPRDGGPSEKSEINEKDLRKK 109
QY 241 SELQGTALGNLKOIYYNSKAITSSSEKADQFLTNLFLPKGFFTHGPWVNDLLVLGSTA 300
DB 110 -----ICTGSEVSS-----VFIFPPKPKDVL----- 130
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPIINLWIDG-- 358
DB 131 -----TITL-----TP-KVTCVVVDIS-----QDDPEVHFSWFVDVVE 162
QY 359 ---KOTTVPIDKVTSSKKEVTQVELDQARHYLHGKFLGFLYNSDSFGGKVGQRLIVFHSSE 415
DB 163 VHTAOTRPEEFGNFTFR--SVSELPILHODWLNGR----- 196
QY 416 GSTVSYDLFDAOGQYPTDLLRIYRDNTTISSTLSISLYLTTSIVMTQTPTSLVLSAGD 475
DB 197 ----- 196
QY 476 RVTITCKASQSVSNDVAVYQOKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDFLTITSS 535
DB 197 ---TFRCKV-----TSAAFPSPIEKTISK 217
QY 536 VOAEADAAYVFCOODYNSPPTFGGGTKLEIKRADAAPTVSIFPPSSEQLTSGGASVYVCFN 595
DB 218 PE-----GRQV-----PHVYTMSPTEEMTQNEVSTCWVK 249
QY 596 NPYPKDINVKWKIDGSRQNGVLSNWTDDQSKDSYSSMSSTLTITKDEYERHNSYTCAT 655
DB 250 GFYPDIIYVQMNGQPQEN-YKNTPPTMDT-DGSFVLYSKLVNKKKQWQGNQGTFTCSVL 307
QY 656 HK 657
DB 308 HE 309

RESULT 8
GCA_RAT
ID GCA_RAT STANDARD; PRT; 322 AA.
AC P20760;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2A chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13804; AAA41376.1; ALT_INIT.
DR PIR; PS0019; PS0019.
DR HSP; P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.

```



```
Db 61 LYTSSSVTVTSWPSSQITCNVAHPASSTKVKKIEPR-----GPTIKPCP----- 108
QY 241 SELOGTALGNLQKIYYNSKAITSSSEKSAQDLTNTLLFKGFTGHPWYNLLVDLGSTA 300
Db 109 ----- 108
QY 301 ATSEYEGSVDLYGAYGYQAGGTPNKTACWYGVTLHDNNRLTEKKVPINLWIDGKQ 360
Db 109 -----PCKCPAPN-----LLGSPSVF----- 124
QY 361 TTVPIDKVKTSKEVTVOELDQARHYLHGKEGLYNSDSFGCKVQORGLIVFHSSEGSTVS 420
Db 125 -----IFFPKIKDVLMI-----SLSPIVT 143
QY 421 YDLFDAQGOYPTLLRIYRDNTTISSTLSISLYLTTSIVMTQPTSLLSVAGDRVTIT 480
Db 144 CVVDVSEDDPD-----VQISWFVNNVEVHTAQQT----- 174
QY 481 CKASOSVNDVAVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTSSVQAE 540
Db 175 -----HREDYNSTLRVV----- 186
QY 541 AAVYFCQDYNSTPTFGGKTLEIKRAD-----AAPTYSIFPPSSEQILT 584
Db 187 SALPTQHQDWMSSGKEF-----CKVNNKDLPAPIERTISKPGSVRAPQVYVLPPEEEMT 242
QY 585 SGGASVVFCLNIFYPKDINVKWKIDGSRQN-----GVLSNWTQDQSDKSTYSMSSTLT 639
Db 243 KQOVLTCMTVDMPEDIYVENNKGKTELVNKNTEPVLDS-----DGSYFMYSKLRV 295
QY 640 TKDEYERHNSYTCETHK-TTSPIVKSFNR 669
Db 296 EKKNWVERNSYSCSVVHEGLHNHHTTKFSR 326

RESULT 10
GCAM_MOUSE
ID GCAM_MOUSE STANDARD; PRT; 399 AA.
AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE I9 gamma-2A chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains."
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -! ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -! MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00471; AAB59661.1; ALT_INIT.
```

```
DR PIR; A02154; G2MSAM.
DR HSSP; P01842; 7FAB.
DR MGD; MG1:96443; Igh-1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363
FT DOMAIN 364 399
FT CARBOHYD 180 180
FT CARBOHYD 180 180
SQ SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;

Query Match 14.0%; Score 491.5; DB 1; Length 399;
Best Local Similarity 26.4%; Pred. No. 2e-23;
Matches 151; Conservative 43; Mismatches 110; Indels 267; Gaps 13;

QY 121 AKTTPSPYPLAPGSAATNSMTLGLCLVKGYFPEPVTVTVNSGSLSSGVTFTPAVLQSD 180
Db 1 AKTTAPSVYPLAPVCGDGTGSSVTLCGLVKGYFPEPVTLTNWNSGSLSSGVTFTPAVLQSD 60
QY 181 LYTSSSVTVSPSTWPSSETVTCNVAHPASSTKVKKIYVPROSGGSPSEKSEENKDLRKK 240
Db 61 LYTSSSVTVSTWPSOSITCNVAHPASSTKVKKIEPR---GPTIKPCP----- 108
QY 241 SELOGTALGNLQKIYYNSKAITSSSEKSAQDLTNTLLFKGFTGHPWYNLLVDLGSTA 300
Db 109 ----- 108
QY 301 ATSEYEGSVDLYGAYGYQAGGTPNKTACWYGVTLHDNNRLTEKKVPINLWIDGKQ 360
Db 109 -----PCKCPAPN-----LLGSPSVF----- 124
QY 361 TTVPIDKVKTSKEVTVOELDQARHYLHGKEGLYNSDSFGCKVQORGLIVFHSSEGSTVS 420
Db 125 -----IFFPKIKDVLMI-----SLSPIVT 143
QY 421 YDLFDAQGOYPTLLRIYRDNTTISSTLSISLYLTTSIVMTQPTSLLSVAGDRVTIT 480
Db 144 CVVDVSEDDPD-----VQISWFVNNVEVHTAQQT----- 174
QY 481 CKASOSVNDVAVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTSSVQAE 540
Db 175 -----HREDYNSTLRVV----- 186
QY 541 AAVYFCQDYNSTPTFGGKTLEIKRAD-----AAPTYSIFPPSSEQILT 584
Db 187 SALPTQHQDWMSSGKEF-----CKVNNKDLPAPIERTISKPGSVRAPQVYVLPPEEEMT 242
QY 585 SGGASVVFCLNIFYPKDINVKWKIDGSRQN-----GVLSNWTQDQSDKSTYSMSSTLT 639
Db 243 KQOVLTCMTVDMPEDIYVENNKGKTELVNKNTEPVLDS-----DGSYFMYSKLRV 295
QY 640 TKDEYERHNSYTCETHK-TTSPIVKSFNR 669
Db 296 EKKNWVERNSYSCSVVHEGLHNHHTTKFSR 326

RESULT 11
HV12_MOUSE
ID HV12_MOUSE STANDARD; PRT; 117 AA.
```

```
AC P01756;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ig heavy chain v region MOPC 104E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE AND CARBOHYDRATE-LINKAGE SITE ASN-55.
RX MEDLINE=83075344; PubMed=6816276;
RA Kehry M.R., Fuhrman J.S., Schilling J.W., Rogers J., Sibley C.H.,
RA Hood L.E.;
RT "Complete amino acid sequence of a mouse mu chain: homology among
RT heavy chain constant region domains.";
RL Biochemistry 21:5415-5424(1982).
CC -!- MISCELLANEOUS: THE SEQUENCE OF THE LIGHT CHAIN OF THIS IGM MYELOMA
CC -!- PROTEIN HAS ALSO BEEN DETERMINED.
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.
DR PIR; A02039; MHMS4E.
DR HSSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin v region; Glycoprotein.
KW DISULFID 22 96 BY SIMILARITY.
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (COMPLEX).
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12983 MW; 3CF8ACE4BE447E41 CRC64;

Query Match 13.7%; Score 481.5; DB 1; Length 117;
Best Local Similarity 78.3%; Pred. No. 1.6e-23;
Matches 94; Conservative 8; Mismatches 15; Indels 3; Gaps 1;

QY 1 EVLOQSGPDLVPGASVKISCKASYSTGYGMHWKQSPGKLEWIGRINPNNGVTL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 EVLOQSGPELVPGASVKMSCKASYSTGYGMHWKQSHGKSLWIGDINPNNGVTS 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NQFKDKATLTVDKSSSTAYMELRSLTSDSAVYICARSTMTINYMDYWGQTSVTSS 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 NQFKDKATLTVDKSSSTAYMQLNSLTSDSAVYICARD---YDWFYDVGAGTTVTSS 117
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
HV07_MOUSE STANDARD; PRT; 139 AA.
AC P01751; P01752;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain v region B1-8/186-2 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234348; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the Npb family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:623-637(1981).
CC -!- MISCELLANEOUS: THE B1-8 MU CHAIN MRNA WAS CLONED FROM A HYBRIDOMA
CC MAKING ANTIBODIES TO THE HAPTEN (4-HYDROXY-3-NITROPHENYL)ACETYL
CC (NPB ANTIBODIES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00529; AAA38170.1; -
DR PIR; A02034; MHMS18.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin v region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 139 IG HEAVY CHAIN V REGION B1-8/186-2.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DOMAIN 118 124 D SEGMENT.
FT DOMAIN 125 139 JH2 SEGMENT.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 139 139
SQ SEQUENCE 139 AA; 15419 MW; 1B57DD4FDC09F465 CRC64;

Query Match 13.7%; Score 481; DB 1; Length 139;
Best Local Similarity 74.2%; Pred. No. 2.1e-23;
Matches 89; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

QY 1 EVLOQSGPDLVPGASVKISCKASYSTGYGMHWKQSPGKLEWIGRINPNNGVTL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 20 QVLOQPGAEVLKPGASVKLSCKASYSTGYGMHWKQSPGKLEWIGRIDPNSGGTKY 79
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NQFKDKATLTVDKSSSTAYMELRSLTSDSAVYICARSTMTINYMDYWGQTSVTSS 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 80 NEKFKSKATLTVDKSPSTAYMQLNSLTSDSAVYICARYDYGGSYFDYWGQGTTLTVSS 139
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
HV13_MOUSE STANDARD; PRT; 117 AA.
AC P01757;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain v region J559.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=80078170; PubMed=6765983;
RA Schilling J., Clevinger B., Davie J.M., Hood L.;
RT "Amino acid sequence of homogeneous antibodies to dextran and DNA
RT rearrangements in heavy chain v-region gene segments.";
RL Nature 283:35-40(1980).
CC -!- MISCELLANEOUS: THE SEQUENCES OF 10 HYBRIDOMA PROTEINS THAT ALSO
CC BIND DEXTRAN DIFFER FROM THAT SHOWN AT 1-7 POSITIONS, MANY OF
CC WHICH OCCUR IN THE D AND J SEGMENTS.
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.
DR PIR; A26242; MHMSJ5.
DR HSSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin v region.
KW DISULFID 22 96 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13024 MW; 292E2AF4BE447E41 CRC64;
```

```

Query Match          13.6%; Score 480.5; DB 1; Length 117;
Best Local Similarity 78.3%; Pred. No. 1.8e-23;
Matches 94; Conservative 7; Mismatches 16; Indels 3; Gaps 1;

QY 1 EVOLQSGPLVPGASVKISKASGYSTGYVMHWKQSPGKLEWIGRINPNNGVTL 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 EVOLQSGPLVPGASVKISKASGYSTGYVMHWKQSPGKLEWIGRINPNNGVTL 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 NQKFKDKATLVKSSSTAYMELRSLTSDSAVYVCARSTMTINYYMDYWGQCTSVTVSS 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 NQKFKDKATLVKSSSTAYMELRSLTSDSAVYVCARSTMTINYYMDYWGQCTSVTVSS 117
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14
HV51_MOUSE          STANDARD;          PRT;          118 AA.
ID  HV51_MOUSE
AC  P06330;
DT  01-JAN-1988 (Rel. 06, Created)
DT  01-JAN-1988 (Rel. 06, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  Ig heavy chain V region AC38 205.12.
OS  Mus musculus (Mouse)
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE.
RX  MEDLINE=84182519; PubMed=6201362;
RA  Dillard R., Bovens J., Stekevitz M., Beyreuther K., Rajewsky K.;
RA  "A V region determinant (idiotope) expressed at high frequency in B
RT  lymphocytes is encoded by a large set of antibody structural genes.";
RL  EMBO J. 3:517-523(1984).
DR  PIR: A02040; MHMS38.
DR  HSSP: P01789; IMCP.
DR  InterPro: IPR003006; Ig_MHC.
DR  InterPro: IPR003596; Ig_V.
DR  Pfam: PF00047; Ig; 1.
DR  SMART: SM00406; IGV; 1.
KW  Immunoglobulin V region.
FT  DOMAIN 1 98 V SEGMENT.
FT  DOMAIN 99 104 D SEGMENT.
FT  DOMAIN 105 118 J SEGMENT.
FT  DISULFID 22 96 BY SIMILARITY.
FT  NON_TER 118 118
SQ  SEQUENCE 118 AA; 12934 MW; 94F7BEE4C762A018 CRC64;

Query Match          13.6%; Score 480; DB 1; Length 118;
Best Local Similarity 79.2%; Pred. No. 2e-23;
Matches 95; Conservative 6; Mismatches 17; Indels 2; Gaps 1;

QY 1 EVOLQSGPLVPGASVKISKASGYSTGYVMHWKQSPGKLEWIGRINPNNGVTL 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 EVOLQSGPLVPGASVKISKASGYSTGYVMHWKQSPGKLEWIGRINPNNGVTL 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 NQKFKDKATLVKSSSTAYMELRSLTSDSAVYVCARSTMTINYYMDYWGQCTSVTVSS 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 NQKFKDKATLVKSSSTAYMELRSLTSDSAVYVCARGYDYPF--DYGWGTGTTVTSS 118
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 15
GCB_MOUSE          STANDARD;          PRT;          336 AA.
ID  GCB_MOUSE
AC  P01866;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Ig gamma-2B chain C region.
OS  Mus musculus (Mouse)
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A. (A ALLELE).

```

```

RX MEDLINE=80120716; PubMed=6766534;
RA Yamawaki-Kataoka Y., Kataoka T., Takahashi N., Obata M., Honjo T.;
RT "Complete nucleotide sequence of immunoglobulin gamma2b chain gene
RN cloned from newborn mouse DNA.";
RL Nature 283:786-789(1980).
RX SEQUENCE FROM N.A. (MPC 11).
RX MEDLINE=80081501; PubMed=117548;
RA Tucker P.W., Marcu K.B., Slightom J.L., Blattner F.R.;
RT "Structure of the constant and 3' untranslated regions of the murine
RN gamma 2b heavy chain messenger RNA.";
RL Science 206:1299-1303(1979).
RX SEQUENCE FROM N.A.
RX MEDLINE=80081502; PubMed=117549;
RA Tucker P.W., Marcu K.B., Newell N., Richards J., Blattner F.R.;
RT "Sequence of the cloned gene for the constant region of murine gamma
RN 2b immunoglobulin heavy chain.";
RL Science 206:1303-1306(1979).
RX SEQUENCE FROM N.A. (B ALLELE).
RX MEDLINE=82173203; PubMed=6803173;
RA Ollo R., Rougeon F.;
RT "Mouse immunoglobulin allotypes: post-duplication divergence of gamma
RN 2a and gamma 2b chain genes.";
RL Nature 296:761-763(1982).
RX CARBOHYDRATE-LINKAGE SITE THR-105.
RX MEDLINE=94216359; PubMed=7512967;
RA Kim H., Yanaguchi Y., Masuda K., Matsunaga C., Yamamoto K.,
RA Irimura T., Takahashi N., Kato K., Arata Y.;
RT "O-glycosylation in hinge region of mouse immunoglobulin G2b.";
RL J. Biol. Chem. 269:12345-12350(1994).
CC -!- PTM: O-LINKED GLYCAN CONSISTS OF GAL-GALNAC DISACCHARIDE WHICH IS
CC MODIFIED WITH 2 SIALIC ACID RESIDUES.
CC -!- MISCELLANEOUS: DISULFIDE BONDS BY HOMOLOGY WITH OTHER IG GAMMA
CC CHAINS.
CC -!- MISCELLANEOUS: THE A ALLELE SEQUENCE IS SHOWN.
DR PIR: A02157; G2MS11.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig-C1.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00407; IGV; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 118 118 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 150 210
FT DISULFID 256 314
FT CARBOHYD 105 105
FT MOD_RES 336 336 REMOVED POST-TRANSLATIONALLY (PROBABLE).
FT VARIANT 163 163 Q -> R (IN B ALLELE).
FT VARIANT 194 194 T -> A (IN B ALLELE).
FT VARIANT 300 300 N -> D (IN B ALLELE).
FT VARIANT 301 301 M -> I (IN B ALLELE).
FT CONFLICT 25 25 L -> S (IN REF. 2 AND 3).
FT CONFLICT 36 36 S -> P (IN REF. 2 AND 3).
FT CONFLICT 239 239 I -> T (IN REF. 2 AND 3).
SQ SEQUENCE 336 AA; 36658 MW; 7D879662607C356E CRC64;

Query Match          13.6%; Score 479; DB 1; Length 336;
Best Local Similarity 27.6%; Pred. No. 9.2e-23;
Matches 151; Conservative 47; Mismatches 110; Indels 240; Gaps 14;

QY 121 AKTTPSPVPLAPGSAQTNSMVTLCGLVKGYFPDPVTWTNWSGSLSGVHTFPAVLQSD 180

```

```
Db 1 AKTTPPSVYPLAPGCGDTGSSVTLGCLVKGYFPEVTVTNWNSGLSSVHTFPALLQSG 60
Qy 181 LYTSSSVTPSSWPSSTVTCNVAHPASSTKVDKIKIVPRDGGPSEKSEKSEINE-----KD 236
Db 61 LYTSSSVTPSSWPSSTVTCNVAHPASSTVDKLEP-----SGP---ISTINPCPKCE 114
Qy 237 LRK--KSELOGTALGNLKOIYYNSKAITSEKSDAQFLTNLTKGFTGHPWVNDLLV 294
Db 115 CHKCPAPNLEG---GPSVFIFPPNIKDV-----LMI 142
Qy 295 DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINL 354
Db 143 SL-----TP----- 146
Qy 355 WIDGKQTTVPIDKVKTSKKEVTVOELDQARHYLHGKFLYNSDSFGKGKVGORGLIVFHSS 414
Db 147 ---KVTGVVD-----VSEDDPDVQISWVFN-----VEVHTA 176
Qy 415 EGSTVSYDLFDAQGOYPTDLLRIYRDNNTISSTLSISLYTTSIVMTQPTSLLSVAG 474
Db 177 QTQTHREDY-----NSTIRVVS----- 193
Qy 475 DRVTITCRASQSVNDVAVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDFLTIS 534
Db 194 -----TLPIQ 198
Qy 535 SVQAEDAAVYFCQDYNPSPTFGGGTKLEIKRADAAPTVSIFPPSPSEQLTSGGASVWCFL 594
Db 199 HDWNSGKFEKCKYNNKDLPSPIERTISKIGLVRAPQVYILPPPAEQLSRKDVSLTCLV 258
Qy 595 NNFYPKDINVKWKIDGSRQN-----GVLNSWTDDSKDSTYSMSSTLTITKDEYERHNS 649
Db 259 VGFNPGLDISVWTSNGHTEENYKDTAPVLDSE-----DGSYFYISKLNKMKTSKWEKTD 311
Qy 650 YTCEATHK 657
Db 312 FSCNVRHE 319
RESULT 16
GCBM_MOUSE
ID GCBM_MOUSE STANDARD: PRT: 405 AA.
AC P01867;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 335-405 FROM N.A.
RX MEDLINE=8222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
RN [2]
RP SEQUENCE OF 335-378 FROM N.A.
RX MEDLINE=8211295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
RT immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
```

```
CC -!- MISCELLANEOUS: DISULFIDE BONDS BY HOMOLOGY WITH OTHER IG GAMMA
CC CHAINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-335 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM.
CC -!- MISCELLANEOUS: THE A ALLELE SEQUENCE IS SHOWN.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00462; AAB59659.1; ALT_INIT.
CC PIR: C02154; G2MSBM.
CC HSSP: P01842; 7FAB.
CC MGI: 96445; Igh-3.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003597; Ig_C1.
CC Pfam: PF00047; Ig; 3.
CC SMART: SM00410; Ig_Like; 1.
CC SMART: SM00407; IG_C1; 2.
CC PROSITE: PS00290; Ig_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Transmembrane;
CC Alternative splicing.
CC NON_TER 1 1 INTERCHAIN (WITH A LIGHT CHAIN).
CC DISULFID 15 15
CC DISULFID 27 82
CC DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 118 118 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 150 210
CC DISULFID 256 314
CC DISULFID 352 369
CC TRANSMEM 370 405
CC DOMAIN 370 405 POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC SEQUENCE 405 AA; 44330 MW; 89B3CF0A9B6D49FA CRC64;
Query Match 13.68; Score 479; DB 1; Length 405;
Best Local Similarity 27.68; Pred. No. 1.2e-22;
Matches 151; Conservative 47; Mismatches 110; Indels 240; Gaps 14;
Qy 121 AKTTPPSVYPLAPGSAQTNSMVTGLCLVKGYFPEVTVTNWNSGLSSVHTFPALLQSD 180
Db 1 AKTTPPSVYPLAPGCGDTGSSVTLGCLVKGYFPEVTVTNWNSGLSSVHTFPALLQSG 60
Qy 181 LYTSSSVTPSSWPSSTVTCNVAHPASSTKVDKIKIVPRDGGPSEKSEKSEINE-----KD 236
Db 61 LYTSSSVTPSSWPSSTVTCNVAHPASSTVDKLEP-----SGP---ISTINPCPKCE 114
Qy 237 LRK--KSELOGTALGNLKOIYYNSKAITSEKSDAQFLTNLTKGFTGHPWVNDLLV 294
Db 115 CHKCPAPNLEG---GPSVFIFPPNIKDV-----LMI 142
Qy 295 DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINL 354
Db 143 SL-----TP----- 146
Qy 355 WIDGKQTTVPIDKVKTSKKEVTVOELDQARHYLHGKFLYNSDSFGKGKVGORGLIVFHSS 414
Db 147 ---KVTGVVD-----VSEDDPDVQISWVFN-----VEVHTA 176
Qy 415 EGSTVSYDLFDAQGOYPTDLLRIYRDNNTISSTLSISLYTTSIVMTQPTSLLSVAG 474
Db 177 QTQTHREDY-----NSTIRVVS----- 193
Qy 475 DRVTITCRASQSVNDVAVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDFLTIS 534
Db 194 -----TLPIQ 198
Qy 535 SVQAEDAAVYFCQDYNPSPTFGGGTKLEIKRADAAPTVSIFPPSPSEQLTSGGASVWCFL 594
```

```
Db 199 HQWMHSGREFKCKVNNKDLPSPIRTISKIRGLVRAPQVILPPPAEQLSKRDVSLTCLV 258
Qy 595 NNFYKIDINVKWKIDGSRQN-----GVLNSWTDDSKDSTYSMSSTLTITKDEYERHNS 649
Db 259 VGFNPGDISVEMTSNGHTEENYKDTAPVLDS-----DGSTFYISKLNMTKTSKWEKTD 311
Qy 650 YTCATHK 657
Db 312 FSCNVRHE 319
RESULT 17
KACA_RAT
ID KACA_RAT STANDARD; PRT; 106 AA.
AC P01836;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain C region, A allele.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA;
RX MEDLINE=82082587; PubMed=6273908;
RA Sheppard H.W., Gutman G.A.;
RT "Allelic forms of rat kappa chain genes: evidence for strong
RT selection at the level of nucleotide sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:7064-7068(1981).
DR PIR: A02118; K1RTA.
DR HSSP: P01842; 2MCG.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00407; IGcl; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 26 86
FT DISULFID 106 106
SQ SEQUENCE 106 AA; 11732 MW; B7E120D9700DD66 CRC64;
Query Match 13.4%; Score 472; DB 1; Length 106;
Best Local Similarity 83.8%; Pred. No. 5.4e-23;
Matches 88; Conservative 11; Mismatches 6; Indels 0; Gaps 0;
Qy 567 ADAAPTTSIFPPSPQLTSGGASVYCFLNFPKIDINVKWKIDGSRQGVLSWTDDSD 626
Db 1 ADAAPTTSIFPPSPQLTSGGATVYCFVNNFYPRDISVKWKIDGSEQRDVLSVTDDSD 60
Qy 627 KDSTYSMSSTLTITKDEYERHNSYTCATHKTSPIVKSFNRRNE 671
Db 61 KDSTYSMSSTLTITKVEYERHNLTYTCEVYVHRTKTSPPVKSFNRRNE 105
RESULT 18
GCAB_MOUSE
ID GCAB_MOUSE STANDARD; PRT; 335 AA.
AC P01864;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2A chain C region, B allele.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
```

```
RX MEDLINE=82037861; PubMed=6170065;
RA Schreier P.H., Bothwell A.L.M., Mueller-Hill B., Baltimore D.;
RT "Multiple differences between the nucleic acid sequences of the
RT IgG2aa and IgG2ab alleles of the mouse";
RL Proc. Natl. Acad. Sci. U.S.A. 78:4495-4499(1981).
RN [2]
RP SEQUENCE.
RX MEDLINE=82037777; PubMed=6794027;
RA Dognin M.J., Lauwereys M., Strosberg A.D.;
RT "Multiple amino acid substitutions between murine gamma 2a heavy
RT chain FC regions of Ig1a and Ig1b allotypic forms.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:4031-4035(1981).
CC -!- MISCELLANEOUS: THE SEQUENCE DIFFERS FROM THAT OF THE A ALLELE,
CC FROM BALB/C MICE, AT 15% OF THE POSITIONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00479; -; NOT_ANNOTATED_CDS.
DR PIR: A02153; G2MSAB.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00407; IGcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
SQ SEQUENCE 335 AA; 36596 MW; FA3382792CBB13C6 CRC64;
Query Match 13.3%; Score 467; DB 1; Length 335;
Best Local Similarity 25.5%; Pred. No. 5.1e-22;
Matches 149; Conservative 40; Mismatches 110; Indels 286; Gaps 13;
Qy 121 AKTTPPSVYPLAPGSAATNSMTVLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
Db 1 AKTAPSVYPLVPVCGGTGTGSSVTLGCLVKGYFPEPVTLTWNSGSLSSGVHTFPAVLQSG 60
Qy 181 LYTLSSTVTPSSTWPSSETVCNVAHPASSTKVDKKIIPRDSGGPSEKSEINEKDLRKK 240
Db 61 LYTLSSTVTPSSTWPSSETVCNVAHPASSTKVDKKIIPR----- 100
Qy 241 SELQGTALGNLKOIYVYNSKAITSSKSAQDQFLNTLLFKGFTTGHWPYNDLLVDLGSTA 300
Db 101 ----- 100
Qy 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACWYGGVTLHDNNRLTEKKVYINLWIDGKQ 360
Db 101 -----VPIT-----QN 106
Qy 361 TTVPIDKVKTSKKEVTYQELDLQARHYLHGKFGLYNSDFGKVGQGLIVFHSSEGSTVS 420
Db 107 PCPPHQVRVPPCAA-----PDLGGP----- 126
Qy 421 YDLFDAQGYQPDTLRLIYRDNTTISSTLSISLYLYTTTSIVMTQTPTSLVLSAGDRVTIT 480
Db 127 -SVIFPPKIKDVLN-----ISLSPWVTCVVV----- 152
Qy 481 CKASQSVSNDAVWYQOKPGQSPKLLIYSTSSRYAGVDPDRFSGSGYGTDTITISVQAED 540
Db 153 -----DVSED-----DPDQVQISW-----FVNNEVHT 174
Qy 541 AAVYFCQDDYNSPTTFGGGTGKLEIKRAD-----A 569
Db 175 AQTQTHREDYNS--TLRVVSALPIQHODMWSGKEFKCKVNNRNLPSPIEKTISKPRGVR 232
```


or send an email to license@isb-sib.ch).

```

CC -----
CC EMBL: J00539; AAA38172.1; -.
CC PIR: A02038; G2MS43.
CC HSSP: P01810; 2FBJ.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003596; Ig_v.
CC Pfam: PF00047; Ig; 1.
CC SMART: SM00406; IGV; 1.
CC Immunoglobulin V region; Signal.
KW SIGNAL 1 19
FT CHAIN 20 137 IG HEAVY CHAIN V REGION S43.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DOMAIN 118 122 D SEGMENT.
FT DOMAIN 123 137 JH2 SEGMENT.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 137 137
SQ SEQUENCE 137 AA; 15200 MW; ADD5881BF44B8EC9 CRC64;

Query Match 12.9%; Score 454; DB 1; Length 137;
Best Local Similarity 70.0%; Pred. No. 9.8e-22;
Matches 84; Conservative 15; Mismatches 19; Indels 2; Gaps 1;

Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYTHHWYKQSPGKGLGWIGRINPNNGVTLY 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 QVQLQQSGAEFLVRGSSVKMSCKASGYTFTSYGNWVKRQPGQGLEWIGYINPGNGYIN 79
Qy 61 NQKFKDKATLVDRKSSSTAYMELRSLSSEDSAVYYCARSTMTI-TNYVMDYWGQGTSTVTS 120
Db :|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 NEHFRSKATLVDRKSSSTAYMELRSLSSEDSAVYYCARVRL--GRYFDYWGQGTTLTVSS 137

RESULT 22
HV02_MOUSE
ID HV02_MOUSE STANDARD; PRT; 140 AA.
AC P01746;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 93c7 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/J;
RX MEDLINE=82152818; PubMed=6801765;
RA Sins J., Rabbits T.H., Estess P., Slaughter C., Tucker P.W.,
RA Capra J.D.;
RT "Somatic mutation in genes for the variable portion of the
RT immunoglobulin heavy chain.";
RL Science 216:309-311(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: J00493; AAA38128.1; -.
CC PIR: A02028; HVM5G7.
CC HSSP: P01810; 2FBJ.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003596; Ig_v.
CC Pfam: PF00047; Ig; 1.
CC SMART: SM00406; IGV; 1.

```

```

KW Immunoglobulin V region; Antiarsenate antibody; Hybridoma; Signal.
FT SIGNAL 1 19
FT CHAIN 20 140 IG HEAVY CHAIN V REGION 93G7.
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 15514 MW; 25A4CBBE31DA5CE8 CRC64;

```

Query Match 12.8%; Score 450.5; DB 1; Length 140;
Best Local Similarity 71.9%; Pred. No. 1.7e-21;
Matches 87; Conservative 14; Mismatches 19; Indels 1; Gaps 1;

```

Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSTGYTHHWYKQSPGKGLGWIGRINPNNGVTLY 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 EVQLQQSGAEFLVRGSSVKMSCKASGYTFTSYGNWVKRQPGQGLEWIGYINPGNGYIN 79
Qy 61 NQKFKDKATLVDRKSSSTAYMELRSLSSEDSAVYYCARSTMTI-TNYVMDYWGQGTSTVTS 119
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 NEKFKGKTTLVDRKSSSTAYMELRSLSSEDSAVYYCARSHYGGSYDFDYWGQGTPLTVS 139
Qy 120 S 120
Db 140 S 140

```

```

RESULT 23
GCB_RAT
ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0018; PS0018.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

```

Query Match 12.7%; Score 446; DB 1; Length 333;
Best Local Similarity 25.3%; Pred. No. 1e-20;
Matches 140; Conservative 54; Mismatches 131; Indels 228; Gaps 14;

```

Qy 121 AKTTPPSVYPLAPGSAQTNSMVTGLCLVKGYFPEPTVTWNSGSLSSGVHTFFAVLQSD 180
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 AQTTPASVYPLAPGCGDFTSTVTGLCLVKGYFPEPTVTWNSGALSSDVHTFFAVLQSG 60
Qy 181 LYTLLSSSVTVPSSTWPSSTVTCNVNAHPASSTKVDKKIVPRDSGGPSEKSEINEKDLRKK 240
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 LYTLLTSSVT--SSTWPSQTVTCNVNAHPASSTKVDKVKERRNG----- 101

```

```
QY 241 SELQ TALGNLKOIYYNKAITSSEKSAQDLNTLLEKGFTHGHPWYNLDLLVGSTA 300
Db 102 -----
QY 301 ATSEYEGSSVDLYGAYGYOCAGGTPNKACMYGGVTLHDNNRLTEEKVPINLWIDGKO 360
Db 102 -----IGHKC-----PTCPTC-----HKCPVPELLGSPS 125
QY 361 TTVPIDKVTSSKEVTVOELDLQARHLHGKFCGLYNSDSFGGKVORGLIVFHSSEGSTVS 420
Db 126 VFI-----FPPKPKDILLI-----SQNAKYT 146
QY 421 YDLFDAQQGYPTDLLRIYRDNNTTISLSLSISLYLTTSIVMTQTPTSLLSAGDRVT-- 478
Db 147 CVVDVSEEPDQVQSFVNNVEVHTAQTPREEQYNSTF---RVVSLPIQHQQHMSCK 203
QY 479 -ITCKASQSVSNDAVWYQKQPGSKLLISYSSRYAGVDPDRFSGSGYGTDTFTLTSSVO 537
Db 204 EFKCKVN-----NKALPSP---IEKTSKPKGLVRK----- 231
QY 538 AEDAAYVFCQDYNPPTPGGKTLEIKRADAAPTYSIFPPSSEQLTSGGASVVCFLNPF 597
Db 232 -----PQYVVMGPPTEQTEQTVSLTCLTSGF 258
QY 598 YPRDINVKKIDGSRQNGVLNSWTDQDSKDSYSSMSSTLTTLTKDEYERHNSYTCETHK 657
Db 259 LPNDIGVETNSGHIEKN-YKNTPEWDS-DGSFFMYSKLNVERSWDSRAPFVCSVVHE 316
QY 658 -TSTSPVKSFRN 659
Db 317 GLHNHHRVKSISR 329

RESULT 24
HV48_MOUSE STANDARD; PRT; 138 AA.
AC P03980;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain v region TEPC 1017 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84248078; PubMed=6429663;
RA Gilliam A.C., Shen A., Richards J.E., Blattner F.R., Mushinski J.F.,
RA Tucker P.W.;
RT "Illegitimate recombination generates a class switch from C mu to C
RT delta in an IgD-secreting plasmacytoma."
RL Proc. Natl. Acad. Sci. U.S.A. 81:4164-4168(1984).
DR PIR; A02033; HVMST7.
DR HSSP; P01810; 2FB3.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_v.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin v region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 138 IG HEAVY CHAIN V REGION TEPC 1017.
FT DOMAIN 21 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DOMAIN 118 127 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 128 138 FRAMEWORK-4.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15576 MW; 748157E4C6907B8E CRC64;
```

```
Query Match 12.4%; Score 437.5; DB 1; Length 138;
Best Local Similarity 68.3%; Pred. No. 1e-20;
Matches 82; Conservative 20; Mismatches 17; Indels 1; Gaps 1;

QY 1 EVLOQSGPDLVKPGASVKISKASGYSTGYIMHWKSPGKLEWIGRIINPNNGVTLY 60
Db 20 QVLOQPGAELVKPGASVOLSKASGHTFTNYIHWKORPGGLEWIGINPNDRSNY 79
QY 61 NQKFKDKATLTVDKSTTAYMELRSLTSEDSAVYYCARSTMTNTNYVMDYWGQSTSTVSS 120
Db 80 NEKFNKATLTVDKSSSTAYMQLSLTPEEFAVYYCARSDGYDWEV-YWGQSTLTFTSA 138

RESULT 25
GCC_RAT STANDARD; PRT; 329 AA.
ID P20762;
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2c chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfrè P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RT region cDNA: extensive homology to mouse gamma 3."
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07189; CAA30169.1; -.
DR PIR; S00847; S00847.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-CL.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IGCL; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 12.4%; Score 437.5; DB 1; Length 329;
Best Local Similarity 26.1%; Pred. No. 3.3e-20;
Matches 140; Conservative 47; Mismatches 125; Indels 225; Gaps 16;

QY 121 AKTTPPSVYPLAPGSAQAQNSMVTGLCLYKGYFPEPVTVTNWNSGSLSSGVHTEPAVLQSD 180
Db 1 ARTTAPSVYPLVPGCGSTSGSLVTLGCLYKGYFPEPVTVTNWNSGALSSGVHTEPAVLQSG 60
```

```
QY 181 LYTSSSVTPSPSTWPTSETVTCNVAHPASTKVYDKKIVPRDGGPSEKSEEINEKDLRKK 240
|||||
Db 61 LYTSSSVTPSPSTWPTSETVTCNVAHPASTKSNLIKRIEPR-----RPKPRPP 107
QY 241 SELQGTALGNLQIYYIYNKAITSSSEKSDQFNTLLFKGFFTHPWNVDLLDLGSTA 300
:::
Db 108 TDT-----CSCDDNLGRPSVF--IFPPKP--KDILM----- 134
QY 301 ATSEYEGSSVDLYGAYGYOCAGGTENKACMGVGTLLHDNNRLTEKKVPGNINLWIDGKQ 360
|||
Db 135 -----ITL-----TP-KVCVVVDVS-----EEPDVQFS-WP----- 160
QY 361 TTVPIDKVKTSKEVTQVQLDQARHLHGKFGCLYNSDSFGKVGQKGLVHFHSSGSETVS 420
|||
Db 161 ----VDNVRVFTAQTPHEEQNLNGTRFVYSTLHIQHODMWSCK----- 199
QY 421 YDLFDAQGOYPTDLLRIYRDNNTISSTLSISLYLYTTSIVMTQTPTSLLVSAGDRVTIT 480
Db 200 -----EKF 202
QY 481 CKASQSVNDVAVYQOKPGSKLLISYTSRVAGVDPDRFSGSGYGTDFTLTSSVQAEQ 540
|||
Db 203 CKVN-----NKDLPSP---TEKTSKPRG----- 223
QY 541 AAVYFCQDYNSPPTFGGKLEIKRADAAPTVSIPTSPSSQELTSGGASVVCFLNNFYPK 600
|||
Db 224 -----KARTPOVYTIPTPPREQMKNKVSILTCWTSFYPA 257
QY 601 DINVKKIDGSRQGVLSWTDQDSKOSTYSMSSTLTLTQDEYERHNSYTCATHK 657
|||
Db 258 SISVEWERNG-ELEQDYKNTLPVLDS-DESYFLYKSLSVDTDSWMRGDIYTCVVVHE 312

RESULT 26
HV50_MOUSE
ID HV50_MOUSE STANDARD; PRT; 120 AA.
AC P06329;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region AC38 15.3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE
RX MEDLINE=84182519; PubMed=6201362;
RA Dildrop R., Bovens J., Silekavitz M., Beyreuther K., Rajewsky K.;
RT "A V region determinant (idiotope) expressed at high frequency in B
RT lymphocytes is encoded by a large set of antibody structural genes.";
RL EMBO J. 3:517-523(1984).
DR PIR; A02037; MHMS15.
DR HSP; P01810; 2FBU.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 186-1.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12890 MW; 16191A088CB17F5A CRC64;

Query Match 12.3%; Score 433; DB 1; Length 120;
Best Local Similarity 67.5%; Pred. No. 1.6e-20;
Matches 81; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

QY 1 EVLOQSGDPLVKPGASVKASVSKSGSYFTGYMHVVKQSPKGLWIGRINPNCVTLY 60
:::
Db 1 EVLOQSGDPLVKPGASVKLSKSGSYFTGYMHVVKQSPKGLWIGRINPNCVTLY 60
QY 61 NQKFKDKATLTVDKSTTAYMELSLTSDSAVYYCAR 98
:::
Db 80 NEKFKSKATLTVDTSSTAYMQLHSLTSDSAVYYCAR 117

RESULT 28
KV5A_MOUSE
ID KV5A_MOUSE STANDARD; PRT; 149 AA.
AC P01633;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region MPC11 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

```
Db 1 QVOLQPGTELKPKGASVNLCKASGYTFTSYMMHWIRQPGGLEWIGINPSNGGTNY 60
QY 61 NQKFKDKATLTVDKSTTAYMELSLTSDSAVYYCARSTMTITNYVMDYWGQTSVTSS 120
:::
Db 61 NEKFKSKATLTVDKSSATYMQSLSTPSTSDSAVYYCARWDYEGDRYFDVWGTTVTSS 120

RESULT 27
HV09_MOUSE
ID HV09_MOUSE STANDARD; PRT; 117 AA.
AC P01753; P11271;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 186-1 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Botwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -I- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
DR PIR; B02034; HVMS61.
DR HSP; P01810; 2FBU.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 186-1.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12890 MW; 16191A088CB17F5A CRC64;

Query Match 12.1%; Score 425; DB 1; Length 117;
Best Local Similarity 79.6%; Pred. No. 4.9e-20;
Matches 78; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 EVLOQSGDPLVKPGASVKASVSKSGSYFTGYMHVVKQSPKGLWIGRINPNCVTLY 60
:::
Db 20 QVOLQPGAEVLKPGASVKLSKSGSYFTSYMMHWIRQPGGLEWIGRINPNCVTLY 60
QY 61 NQKFKDKATLTVDKSTTAYMELSLTSDSAVYYCAR 98
:::
Db 80 NEKFKSKATLTVDTSSTAYMQLHSLTSDSAVYYCAR 117

RESULT 28
KV5A_MOUSE
ID KV5A_MOUSE STANDARD; PRT; 149 AA.
AC P01633;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region MPC11 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```


CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: J00494; AAA38130.1; -.
DR PIR: A02042; HVMSB1.
DR HSP: P01772; 2FB4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig: 1.
DR SMART: SM00406; Igv: 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 136 IG HEAVY CHAIN V REGION BCL1.
FT NON_TER 136 136
FT SEQUENCE 136 AA: 15078 MW: 6827CFBC6DB3F35E CRC64;

Query Match 11.9%; Score 420.5; DB 1; Length 136;
Best Local Similarity 70.0%; Pred. No. 1.1e-19;
Matches 84; Conservative 11; Mismatches 22; Indels 3; Gaps 2;

QY 1 EVOLQSGDPLVKPGASVKISKASGYSTGYMHVWKSPGKLEWIGRINPNNGVTLY 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 20 QVLOQSGPEVVRPGSVKISCKSGYTFDYAMHVKQSHAKSLEWIGVISTYNGTSY 79
QY 61 NOKFKDKATLTVDKSSSTAYMELSLTSDSAVYYCARSTMTNMYDYGOGTSTVSS 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 80 NOKFKDKATMTVDKSSSTVHMLARLTSDSANLYCAR--YYGNY-FDYWGOGTTLTVSS 136

RESULT 31
HV04_MOUSE
ID HV04_MOUSE STANDARD; PRT; 117 AA.
AC P01748;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 23 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the Npb family of
RT antibodies; somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -!- MISCELLANEOUS: THIS GERM LINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.

DR PIR: A02030; HVMS23.
DR HSP: P01810; 2FBJ.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig: 1.
DR SMART: SM00406; Igv: 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19

FT CHAIN 20 117 IG HEAVY CHAIN V REGION 23.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
FT SEQUENCE 117 AA: 12772 MW: C530F829C906F69B CRC64;

Query Match 11.8%; Score 417; DB 1; Length 117;
Best Local Similarity 78.6%; Pred. No. 1.5e-19;
Matches 77; Conservative 9; Mismatches 12; Indels 0; Gaps 0;
QY 1 EVOLQSGDPLVKPGASVKISKASGYSTGYMHVWKSPGKLEWIGRINPNNGVTLY 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 20 QVLOQPGCTELVKPGASVKLSKASGYTFTSYMHVWKQPGQGLEWIGRINPNSGTTY 79
QY 61 NOKFKDKATLTVDKSSSTAYMELSLTSDSAVYYCAR 98
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 80 NEKFKSKVTLTVDKSSSTAYTQLSLTSDSAVYYCAR 117

RESULT 32
HV10_MOUSE
ID HV10_MOUSE STANDARD; PRT; 117 AA.
AC P01754; P11270;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 145 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the Npb family of
RT antibodies; somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -!- MISCELLANEOUS: THIS GERM LINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: J00533; AAA38602.1; -.
DR PIR: C02034; HVMS45.
DR HSP: P01810; 2FBJ.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig: 1.
DR SMART: SM00406; Igv: 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19

FT CHAIN 20 117 IG HEAVY CHAIN V REGION 145.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT NON_TER 117 117
FT SEQUENCE 117 AA: 12921 MW: D37DE8A3F543E996 CRC64;

Query Match 11.8%; Score 415; DB 1; Length 117;
Best Local Similarity 78.6%; Pred. No. 2.1e-19;
Matches 77; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

QY 1 EVOLQSGDPLVKPGASVKISKASGYSTGYMHVWKSPGKLEWIGRINPNNGVTLY 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 20 QVLOQPGRELVKPGASVKLSKASGYTFTSYMHVWKQPGQGLEWIGRINPNSGTTY 79
QY 61 NOKFKDKATLTVDKSSSTAYMELSLTSDSAVYYCAR 98
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```
Db 80 NEFKSKATLTVDKPSSSTAYMOLSLTSEDSAVYYCAR 117
RESULT 33
HV49_MOUSE STANDARD; PRT; 117 AA.
AC P06328;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region VH558 B4 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099340; PubMed=2578321;
RA Yancopoulos G.D., Alt F.W.;
RT "Developmentally controlled and tissue-specific expression of
RT rearranged VH gene segments.";
RL Cell 40:271-281(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13788; AAA38506.1; -.
DR PIR; A02035; MHMSB4.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region; Signal.
KW SIGNAL
FT CHAIN 1 19
FT DOMAIN 20 117 IG HEAVY CHAIN V REGION VH558 B4.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12834 MW; B8862FAC67ABD345 CRC64;

Query Match 11.8%; Score 414; DB 1; Length 117;
Best Local Similarity 78.4%; Pred. No. 2.4e-19;
Matches 76; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

Qy 2 VOLQSGDPLVKPGASVKISCKASGYSTFGYYMHVWKSPGKLEWIGRINPNNGVTLYN 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 21 VOLQPGAEELVKPGASVKLSCKASGYSTFTSYMHVWKQRPGRGLEWIGNIDPNSGRTKN 80
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy 62 QKFKDKATLTVDKSSTAYMELRSLTSEDSAVYYCAR 98
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 81 ERFKSKATLTVDKPSSSTAYMQLSLTSEDSAVYYCTR 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 34
HV06_MOUSE STANDARD; PRT; 117 AA.
ID HV05_MOUSE
AC P01750;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 102 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPb family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -!- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPb ANTIBODIES.
CC PIR; A02032; HVMS02.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region; Signal.
KW SIGNAL
FT CHAIN 1 19
FT DOMAIN 20 117 IG HEAVY CHAIN V REGION 102.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12867 MW; 740A65DD851FCA8C CRC64;

Query Match 11.7%; Score 413; DB 1; Length 117;
Best Local Similarity 80.2%; Pred. No. 2.7e-19;
Matches 77; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

Qy 2 VOLQSGDPLVKPGASVKISCKASGYSTFGYYMHVWKSPGKLEWIGRINPNNGVTLYN 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 21 VOLQPGAEELVKPGASVKVSKASGYSTFTSYMHVWKQRPGRGLEWIGRHPDSDTNYN 80
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy 62 QKFKDKATLTVDKSSTAYMELRSLTSEDSAVYYCA 97
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 81 QKFKDKATLTVDKSSSTAYMQLSLTSEDSAVYYCA 116
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 35
GC2_CAVPO STANDARD; PRT; 329 AA.
ID GC2_CAVPO
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2 chain C region.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
NCBI_TaxID=10141;
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-68.
RX MEDLINE=71058471; PubMed=5538606;
RA Birshstein B.K., Hussain Q.Z., Cebrá J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the
RT half-cystine joining heavy and light chains.";
RL Biochemistry 10:18-25(1971).
RN [3]
RP SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J., Cebrá J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
```



```

Db 2 VLQTFWISLLWISAGYDIVMTQSPDSLAVSLGRATINCKSSQSVLYSSNNKNYLAW 61
      : | : ||| | : : ||||| : | | | | | | | | | | | | | | | | : ||
QY 494 YQKPGQSPKLLISYTSRYAGVDPFRFSGSGYGTDTFTLTSSVQAEADAAYFCQDYNP 553
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |
Db 62 YQKPGQSPKLLIYWASTRESGVDPFRFSGSGTDTFTLTSSVQAEADVAVYQYINLP 120
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |
QY 554 PTFGGGKLEIKR 566
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |
Db 121 PTFGGGKLEIKR 133
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |

RESULT 37
KV4C_HUMAN
ID KV4C_HUMAN STANDARD; PRT: 134 AA.
AC P06314;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region B17 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86041854; PubMed=2997713;
RA Marsh P., Mills F., Gould H.;
RT "Detection of a unique human V kappa IV germline gene by a cloned
RT cDNA probe.";
RL Nucleic Acids Res. 13:6531-6544(1985).
RN [2]
RP REVISION TO 76.
RA Marsh P.;
RL Submitted (OCT-1986) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02990; CAA26733.1; -
DR PIR; A01905; K4HU17.
DR HSP; P80362; LWL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 134 IG KAPPA CHAIN V-IV REGION B17.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 61 75 FRAMEWORK-2.
FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 83 114 FRAMEWORK-3.
FT DOMAIN 115 121 FRAMEWORK-4.
FT DOMAIN 122 133 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 114 BY SIMILARITY.
FT NON_TER 134 134
SQ SEQUENCE 134 AA; 14966 MW; 6413A22FD0738832 CRC64;

Query Match 11.5%; Score 406; DB 1; Length 134;
Best Local Similarity 63.2%; Pred. No. 8.9e-19;
Matches 84; Conservative 15; Mismatches 24; Indels 10; Gaps 2;

QY 444 ISSTLSISLYLT-----SIWMTQPTSLVSGAGRVITCKASQSV-----SNDVAW 493
      : | : ||| | : : ||||| : | | | | | | | | | | | | | | | | : ||
Db 2 VLQTFWISLLWISAGYDIVMTQSPDSLAVSLGRATINCKSSQSVLYSSNNKNYLAW 61
```

```

QY 494 YQKPGQSPKLLISYTSRYAGVDPFRFSGSGYGTDTFTLTSSVQAEADAAYFCQDYNP 553
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |
Db 62 YQKPGQSPKLLIYWASTRESGVDPFRFSGSGTDTFTLTSSVQAEADVAVYQYINLP 121
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |
QY 554 PTFGGGKLEIKR 566
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |
Db 122 WTFGQGTKEIKR 134
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |

RESULT 38
HV05_MOUSE
ID HV05_MOUSE STANDARD; PRT: 117 AA.
AC P01749;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 3 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NpB family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:623-637(1981).
CC -!- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00536; AAA38605.1; -
DR PIR; A02031; HVMSJ.
DR HSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 3.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13016 MW; 427C861C53975EDC CRC64;

Query Match 11.5%; Score 404; DB 1; Length 117;
Best Local Similarity 76.5%; Pred. No. 9.8e-19;
Matches 75; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 EVOLQSGPDLVKPGASVKISCKASGYFTGYMHVVKQSPGKLEWIGRINPNVTLY 60
      : ||||| : | : | | | | | | | | | | | | | | | | | | | |
Db 20 QVLOQPGAEVLPRGSSVLSCKASGYTFTSYMDWVKQRPQGGLEWIGNIYPSDSETH 79
      : ||||| : | : | | | | | | | | | | | | | | | | | | | |
QY 61 NQKFKDKATLTVDKSLTAYMELSLTSEDSAVYYCAR 98
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |
Db 80 NQKFKDKATLTVDKSLTAYMELSLTSEDSAVYYCAR 117
      ||||| | | | | : | : | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 39
GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; ; NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR HSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IG_C1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 11.4%; Score 401; DB 1; Length 329;
Best Local Similarity 24.6%; Pred. No. 6e-18;
Matches 135; Conservative 51; Mismatches 111; Indels 252; Gaps 15;

QY 123 TTPPSVYPLAPGSAATNSMVLGCLVKGYFPPVPTVTVNNSLSGVITFPAVLQSDLY 182
|| ||||| || : ||||| ||||| || ||||| || ||||| || ||||| ||
Db 2 TTPASVYPLVPGSDTSVSGSVTLGCLVKGYFPEPVTVKWNIGALSSGVTSVLSQSGFY 61

QY 183 TLSSSVTPSSVTPSTVTCNVAHPASSFKVDKKIVPRDSGGSPSEKNEKDLRKKSE 242
||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 62 SLSSLVTVFSSVTPSTVTCNVAHPASSFKVDKKIVPRDSGGSPSEKNEKDLRKKSE 105
||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 243 LGCTAL--GNLKOIYYNSKAITSSSEKSAQDLTNTLLEKFGFTGHPWYNLLVDLGSTA 300
|| : || : || : || : || : || : || : || : || : || : || : ||
Db 106 PGSSCPPGNI-----PPK-----LG--- 118

QY 301 ATSEYEGSSVDLYGAYGYQCAGGTGNKACMYGGVTLHDNNRLTEEKVPTNLWDGKQ 360
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 119 -----GPSVFIP-----PPK-----PKDALMSLTPKV 141

QY 361 TTVPIDKVTSSKEVTVQELDLQARHYLHCKFGLYNSDSFGGKVGQRLIVFHSSESTVS 420
|| : || : || : || : || : || : || : || : || : || : || : ||
Db 142 TCVVVD-----VSEDD----- 152

QY 421 YDLFDAGQGYPTDLLRIYRDNTTISSTLSISLYLTTSIVMTQTPTSLVSAGDRVTIT 480
|| : || : || : || : || : || : || : || : || : || : || : ||

```

```

Db 153 -----PDVHVSWEVDNKEVHT----- 168
QY 481 CKASQSVNDVAWYQKQPCQSPKLLISYTSRYAGVDPDRFSSGGYGTDF-TLTISSVQAE 539
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 169 -----AWTQPREAQ-----YNSTRFVVSALPIQHQ 193
QY 540 D---AAVYFCQDYNSPPTFGGKLEIKRADAAPTVSIFPPSSEOLTSGGASVVCFLAN 596
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 194 DMWRGKEFKCKVNNKALPAPIERTISKRGRAQTPQVYTIPTPPREOMSKKKVSLTCLVTN 253
QY 597 FYPKIDINVKWKIDGSRQNGVLNSWTDQDSK-----DSTVSMSTLTLTKDEYERHN 648
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 254 FFESEALSVEW-----RNGEL-----EQDYKNTPTILDSDGYFLYSLKLTVDTDSWLQGE 303
QY 649 SYTCEATHK 657
Db 304 IFTCSVYHE 312

RESULT 40
GC3M_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; AAB59655.1; -
DR EMBL; V01526; CAA24767.1; ALT_SEQ.
DR PIR; A02155; G3MSM.
DR HSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IG_C1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.

```

```
FT TRANSMEM 346 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; C7F264B50A41B95 CRC64;

Query Match 11.4%; Score 401; DB 1; Length 398;
Best Local Similarity 24.6%; Pred. No. 7.7e-18;
Matches 135; Conservative 51; Mismatches 111; Indels 252; Gaps 15;

Qy 123 TTPSPVPLAPCSAQTNSWTLGCLVGYPPPEPVTVTWNSGSLSSGVHTTTPAVLQSDLY 182
  || ||||| || : : ||||| ||||| || : ||||| || : ||||| || : ||||| ||
Db 2 TTPSPVPLPGCSDTSGSVTLGCLVGYPPPEPVTVKWNYGALSSGVRTVSSVLQSGFY 61
  ||||| ||||| || : ||||| ||||| || : ||||| ||||| || : ||||| ||||| ||
Qy 183 TLSSVTVPSSTWPEVTTCNVAHPASSTKVDKLVPRDGGPSEKSEINEKDKLRKKE 242
  ||||| ||||| ||||| || : ||||| ||||| || : ||||| ||||| || : ||||| ||
Db 62 SLSSLVTVPSSTWPEVTTCNVAHPASSTKELIKREPR-----IPKPEST 105
  ||||| ||||| ||||| || : ||||| ||||| || : ||||| ||||| || : ||||| ||
Qy 243 LGTAL--GNLKQIYYNKAITSSEKSAQDQFLTWTLFKGFTTGHWPYNDLLVDLGSTA 300
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 106 PRGSSCPPGNI-----LG--- 118
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Qy 301 ATSEVEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEKVPINLWIDGQ 360
  || || : || : || : || : || : || : || : || : || : || : || : ||
Db 119 -----GPSVFIF-----PPK-----PKDALMISLTPKV 141
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Qy 361 TTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQVGLIVFHSSEGSTVS 420
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 142 TCVVVD-----VSEDD----- 152
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Qy 421 YLFDQAQGYPTLLRIYRDNTTISTSLSTSLYLYTTSIVMTQPTPTLLVSAGDRVTIT 480
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 153 -----PDVHVSWFVDNKEVHT----- 168
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Qy 481 CKASVSNDAVYQOKFGQCKPKLLISTSSRYAGVPRFSGSGYGTDF-TLTISVVAE 539
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 169 -----AMTQPREAQ-----YNSTFRVVSALPIQHQ 193
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Qy 540 D---AAVYFCQDYNSTPFTGGGTGLEIKRADAAPTSTVFPPSSQLTSGGASVVCFLNN 596
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 194 DMWRGKEFKCKVNNKALFAPERTISKPGRAQTPOVTIPPRQMKKKVSLTCLVTN 253
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Qy 597 FYPKDINVKWKIDGSRQNGVLSNSTDQSK-----DSTYSMSSTLTITKDEYERN 648
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 254 FFEAISVEWE-----RNGEL-----EQDYKNTPPILDSGTGYFLYSKLTVDTSWLQCE 303
  || : || : || : || : || : || : || : || : || : || : || : || : ||
Qy 649 SYTCEATHK 657
  :|| : || : || : || : || : || : || : || : || : || : || : || : ||
Db 304 IFTCSVVHE 312
  :|| : || : || : || : || : || : || : || : || : || : || : || : ||

RESULT 41
GC1_HUMAN
ID GC1_HUMAN STANDARD; PRT; 330 AA.
AC P01857;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-1 chain C region.
GN IGHG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82274238; PubMed=6287432;
RA Ellison J.W., Berson B.J., Hood L.E.;
RT "The nucleotide sequence of a human immunoglobulin C gamma1 gene.";
RL Nucleic Acids Res. 10:4071-4079(1982).
RN [2]
RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
-----
```

```
RX MEDLINE=71064024; PubMed=5489771;
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
RA Waxdal M.J., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino
RT acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";
RL Biochemistry 9:3161-3170(1970).
RN [3]
RP SEQUENCE OF 136-329 (EU).
RX MEDLINE=71064025; PubMed=5530842;
RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,
RA Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino
RT acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";
RL Biochemistry 9:3171-3181(1970).
RN [4]
RP SEQUENCE (MYELOMA PROTEIN NIE).
RX MEDLINE=77070269; PubMed=826475;
RA Ponstingl H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a
RT monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The
RT chymotryptic peptides of the H-chain, alignment of the tryptic
RT peptides and discussion of the complete structure.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [5]
RP SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71064027; PubMed=4923144;
RA Gall W.E., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. X.
RT Intrachain disulfide bonds.";
RL Biochemistry 9:3188-3196(1970).
RN [7]
RP DISULFIDE BONDS.
RX MEDLINE=77070267; PubMed=1002129;
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT "Rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein NIE). I: Purification and
RT characterization of the protein, the L- and H-chains, the
RT cyanogen bromide cleavage products, and the disulfide bridges.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=81208100; PubMed=7236608;
RA Deisenhofer J.;
RT "Crystallographic refinement and atomic models of a human FC fragment
RT and its complex with fragment B of protein A from Staphylococcus
RT aureus at 2.9- and 2.8-A resolution.";
RL Biochemistry 20:2361-2370(1981).
CC CC -!- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE
CC CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)
CC CC MARKER & THE GIM (NON-1) MARKERS.
CC CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
CC CC 35, 116, 198, 269 & 272.
CC CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
CC CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
CC CC 268-272.
CC CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
CC CC RESIDUES 198, 267&272.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```



```
Query Match 11.3%; Score 396.5; DB 1; Length 121;
Best Local Similarity 62.8%; Pred. No. 3e-18;
Matches 76; Conservative 16; Mismatches 28; Indels 1; Gaps 1;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYSFTGYMHVWVKSPGKLEWIGRINPNNGVTLY 60
DB 1 EAQLQSGAELVRPGYSVSKISKAAGYFTNYWGWKRPGLHGLWIGDIYPGGGFTNY 60
QY 61 NOKFKDKATLTVDKSSSTAYMELRSLSLSDSAVYICARSTMI--TNVYMDYWGQGTSTVTVS 119
DB 61 NDLKAGKATLTADTSSSTAYIQLSLTSDSAIYHCARGIYVNSSPYFDSWGQGTTLTVS 120
QY 120 S 120
DB 121 S 121

RESULT 43
GC4_HUMAN STANDARD; PRT; 327 AA.
AC P01861;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-4 chain C region.
GN IGHG4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83157104; PubMed=6299662;
RA Ellison J.W., Buxbaum J.N., Hood L.E.;
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
RL DNA 1:11-18(1981).
RN [2]
RP SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE=70207560; PubMed=4192699;
RA Pink J.R.L., Buttery S.H., de Vries G.M., Millstein C.;
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
constant region of a gamma 4 chain.";
RL Biochem. J. 117:33-47(1970).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01316; AAB59394.1; ALT_INIT.
DR PIR; A02150; G4HU.
DR HSP; P01842; 7FAB.
DR Genew; HGNC:5528; IGHG4.
DR MIM; 147130; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 98 CH1..
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 220 CH2.
FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
```

```
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EBDB811EF208E7A CRC64;

Query Match 11.3%; Score 396.5; DB 1; Length 327;
Best Local Similarity 23.3%; Pred. No. 1.1e-17;
Matches 131; Conservative 52; Mismatches 102; Indels 277; Gaps 16;

QY 121 AKTTPPSVYPLAPGSAOQNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSD 180
DB 1 ASTKGPSVPLAPCSKSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS 60
QY 181 -LYTSSSVTVPSSTWPSSTVCNVAHPASSTKVKDKI-----VPRDSGGPSE 227
DB 61 GLYSLSVVTVPSSSLGTYTCNVDRKPSNTKYDRVESKYGPPCPAPAEFLGCP- 119
QY 228 KSEEINEKDLRKKSLOQTALGNLKOIYYNYSKAITSEKSDAQFLTNTLLFKGFTTGH 287
DB 120 -----VFLPPPKP-----KDTLMI----- 133
QY 288 WYNDLLVLDLGSTAATSEYEGSSVDLYGAYGYQCAGGTNKTACMYGGVTLHDNNRLTEE 347
DB 134 -----SRTPEVT-CVVVDVS-----QED 150
QY 348 KKVPINLWIDGKQTTVPIDKVKTSKEV-----TVQELDLQARHYLHGKFLYNSDSF 400
DB 151 PEVQFNWYDVG---VEVHNATKPREQFNSTYRVVSVLTVLHQDLNGLK----- 197
QY 401 GGVQRLIVFHSSEGSTVSYDLDAOGQYPTDLLRIYRDNTTISSTLSISLYLTTSI 460
DB 198 -----EY----- 199
QY 461 VMTQPTSLVLSAGDRVTITCKASQSVNDVAWYQOKPGQSPKLLISYTSRYAGVPDRF 520
DB 200 -----KCKVSNK-----GLP--- 209
QY 521 SGGYGTDFTLTISVQAEAAVYFCQDYNPSPTFGGCTKLEIKRADAAPTYSIFPPSS 580
DB 210 -----SSIEK-----TISKAKQPREPQVYTLPPSQ 235
QY 581 EOLTSGGASVQFLNFPKIDINVKWKIDGSRQN-----GVLNSWTDQSDKSDSYSMSS 635
DB 236 EEMTKNOVSLTCLVKGYFSPDIAVEWESQNPENNYKTTTPVLD- - - - -DGSFELY 288
QY 636 TLTTLTKDEYERHNSYTCEATHK 657
DB 289 RLTVDKSRWQEGNVFSCSYMHE 310

RESULT 44
KV4A_HUMAN STANDARD; PRT; 114 AA.
AC P01625;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region Len.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76004342; PubMed=50995;
RA Schneider M., Hilschmann N.;
RT "The primary structure of a monoclonic immunoglobulin-L-chain of
RT subgroup IV of the kappa type (Bence-Jones protein Len).";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:507-557(1975).
RN [2]
RP REVISION TO 9.
RA Salomon A.;
```

RL Submitted (AUG-1996) to the SWISS-PROT data bank.
CC -!- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC -!- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR PIR: A01903; K4HULN.
DR HSP: P01607; IREI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DOMAIN 1 23
FT DOMAIN 24 40
FT DOMAIN 41 55
FT DOMAIN 56 62
FT DOMAIN 63 94
FT DOMAIN 95 101
FT DOMAIN 102 113
FT DISULFID 23 94
FT NON_TER 114 114
SQ SEQUENCE 114 AA; 12640 MW; 0647F1D17F236485 CRC64;

Query Match 11.2%; Score 395; DB 1; Length 114;
Best Local Similarity 69.9%; Pred. No. 3.4e-18;
Matches 79; Conservative 12; Mismatches 16; Indels 6; Gaps 1;

QY 460 IVMTQPTSLLSVAGDRVTITCKASQSV-----SNDVAWYQKPGQSPKLLISYSSRY 513
DB 2 IVMTQSPDSLAVSLGERATINCKSSQSVLSNKNYLAWYQKPGQPKLLIYNASTRE 61

QY 514 AGVPDRFSGSGYGTDTFTLTISVQAEDAAVYFCQDYNPPTFGGQTKLEIKR 566
DB 62 SGVPDRFSGSGGTDTFTLTISLQAEDVAVYCYQYSTPYSFGQGTKEIKR 114

RESULT 45
KV5B_MOUSE
ID KV5B_MOUSE STANDARD; PRT; 136 AA.
AC P01634;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region MOPC 21 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82059477; PubMed=6170937;
RA Hamlyn P.H., Gait M.J., Milstein C.;
RT "Complete sequence of an immunoglobulin mRNA using specific priming
and the dideoxynucleotide method of RNA sequencing.";
RL Nucleic Acids Res. 9:4485-4494 (1981).
RN [2]
RP SEQUENCE OF 30-136.
RX MEDLINE=73053310; PubMed=4638343;
RA Svasti J., Milstein C.;
RT "The complete amino acid sequence of a mouse kappa light chain.";
RL Biochem. J. 128:427-444 (1972).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: V00810; CAA24192.1; ALT_TERM.
DR PIR: A01917; KVM521.
DR HSP: P80362; IWIU.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.

DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 29
FT CHAIN 30 136
FT DOMAIN 30 52
FT DOMAIN 53 63
FT DOMAIN 64 78
FT DOMAIN 79 85
FT DOMAIN 86 117
FT DOMAIN 118 126
FT DOMAIN 127 136
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 14902 MW; 8CDD85113996D1C2 CRC64;

Query Match 11.2%; Score 395; DB 1; Length 136;
Best Local Similarity 60.9%; Pred. No. 4.3e-18;
Matches 81; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

QY 433 TLLRIYRDNTTISSTLSISLYTTIVMTQPTSLLSVAGDRVTITCKASQSVNDVA 492
DB 4 TSMGKIMESHVTLVIFSLICLCGACGNIVMTQSPKSMVSGERVTLTCKASENVVTVYS 63

QY 493 WYQKPGQSPKLLISYSSRYAGVPDRFSGSGYGTDTFTLTISVQAEDAAVYFCQDYN 552
DB 64 WYQKPEQSPKLLIYGASNRNRYGVPDRFTGSGSATDFTLTISVQAEDLADYHCGGY 123

QY 553 PTFGGGTKEIKR 565
DB 124 PYTFGGGTKEIKR 136

RESULT 46
HV52_MOUSE
ID HV52_MOUSE STANDARD; PRT; 117 AA.
AC P06327;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region VH558 A1/A4 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099340; PubMed=2578321;
RA Yancopoulos G.D., Alt F.W.;
RT "Developmentally controlled and tissue-specific expression of
unrearranged VH gene segments.";
RL Cell 40:271-281 (1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M13787; AAA38499.1; -.
DR PIR: A02029; HVMSA1.
DR HSP: P01810; 2FBI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117
FT DOMAIN 20 49
FT DOMAIN 50 54
FT DOMAIN 55 68

```
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12971 MW; 8B0BC138856DFC9D CRC64;

Query Match 11.1%; Score 392; DB 1; Length 117;
Best Local Similarity 75.5%; Pred. No. 5.4e-18;
Matches 74; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

Oy 1 EVOLQSGPDLVKPCASVKISKASGYSTGYMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 QVQLQSGPELVKQALVKISKASGYFTTSDINWVKRPGQGLEWIGWYPGDGSTKY 79

Oy 61 NOKFKDKATLVKDSSTTAYMELRSLTSEDSAVVYCAR 98
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 NEKFKGKATLTADKSSSTAYMQLSLSLTSENSAVVFCAR 117

RESULT 47
HV1C_HUMAN STANDARD; PRT; 147 AA.
AC P01744;
DT 21-JUL-1986 (Rel. 01, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Ig heavy chain V-I region ND precursor (Fragments).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=83065234; PubMed=6815656;
RA Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,
RA Bell L.O., Gould H.J.;
RT "Cloning and sequence determination of the gene for the human
RT immunoglobulin epsilon chain expressed in a myeloma cell line.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).
RN [2]
RP SEQUENCE OF 20-147.
RA Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;
RL (In) Bach M.K. (eds.);
RL Immediate hypersensitivity: modern concepts and developments, pp.1-36,
RL Marcel Dekker, New York (1978).
CC -I- MISCELLANEOUS: THIS EPSILON CHAIN WAS ISOLATED FROM A MYELOMA
CC PROTEIN.
DR PIR; A02026; ELHUND.
DR HSSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 147 IG HEAVY CHAIN V-I REGION ND.
FT MOD_RES 20 20 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 41 115
FT CONFLICT 21 21 T -> V (IN REF. 2).
FT CONFLICT 53 54 IH -> HI (IN REF. 2).
FT CONFLICT 67 68 VG -> GV (IN REF. 2).
FT CONFLICT 125 125 MISSING (IN REF. 2).
FT NON_TER 147 147
SQ SEQUENCE 147 AA; 16491 MW; 948F9F72A5366C20 CRC64;

Query Match 10.9%; Score 385; DB 1; Length 147;
Best Local Similarity 55.5%; Pred. No. 2e-17;
Matches 71; Conservative 24; Mismatches 25; Indels 8; Gaps 1;

Oy 1 EVOLQSGPDLVKPCASVKISKASGYSTGYMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
20 QTQLVQSGAEVRKPGASRVSKASGYFTTSDYTHWIRQAPGHGLEWYGNPNSSGNTY 79
```

```
Oy 61 NOKFKDKATLVKDSSTTAYMELRSLTSEDSAVVYCAR-----TMTNRYMVDYNGQ 112
Db :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
80 APRFQGRVMTTRDASFSTAYMDLSRLSDSDSAVFYCAKSDPFWSDYNFDSYTLDDVNGQ 139

Oy 113 GTSVTVSS 120
Db ||:|||||
140 GTTVTVSS 147

RESULT 48
GC2_HUMAN STANDARD; PRT; 326 AA.
ID AC P01859;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-2 chain C region.
GN IGHG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=82197621; PubMed=6804948;
RA Ellison J.W., Hood L.E.;
RT "Linkage and sequence homology of two human immunoglobulin gamma
RT heavy chain constant region genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).
RN [2]
RP SEQUENCE OF 88-115 FROM N.A.
RC TISSUE=Fetal liver;
RX MEDLINE=83001943; PubMed=6811139;
RA Takahashi N., Ueda S., Obata M., Nikaide T., Nakai S., Honjo T.;
RT "Structure of human immunoglobulin gamma genes: Implications for
RT evolution of a gene family.";
RL Cell 29:671-679(1982).
RN [3]
RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.
RC TISSUE=Fetal liver;
RX MEDLINE=84235992; PubMed=6329676;
RA Krawinkel U., Rabbitts T.H.;
RT "Comparison of the hinge-coding segments in human immunoglobulin gamma
RT heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass
RT genes.";
RL EMBO J. 1:403-407(1982).
RN [4]
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
RX MEDLINE=81007873; PubMed=6774012;
RA Wang A.-C., Tung E., Fudenberg H.H.;
RT "The primary structure of a human IgG2 heavy chain: genetic,
RT evolutionary, and functional implications.";
RL J. Immunol. 125:1048-1054(1980).
RN [5]
RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).
RX MEDLINE=80001357; PubMed=113060;
RA Connell G.E., Parr D.M., Hofmann T.;
RT "The amino acid sequences of the three heavy chain constant region
RT domains of a human IgG2 myeloma protein.";
RL Can. J. Biochem. 57:758-767(1979).
RN [6]
RP SEQUENCE OF 238-275 (ZIE).
RX MEDLINE=80114419; PubMed=118920;
RA Hofmann T., Parr D.M.;
RT "A note of the amino acid sequence of residues 381-391 of human
RT immunoglobulins gamma chains.";
RL Mol. Immunol. 16:923-925(1979).
RN [7]
RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).
RA Hofmann T., Parr D.M.;
RL Submitted (MAR-1980) to the PIR data bank.
RN [8]
RP SEQUENCE OF 1-121 (DOT).
```

RX MEDLINE=9525298; PubMed=7737190;
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
 RT "Characterization of the two unique human anti-flavon monoclinal
 RL immunoglobulins";
 RN Eur. J. Biochem. 228:886-893(1995).
 RP [9]
 RP DISULFIDE BONDS.
 RX MEDLINE=72033500; PubMed=4940472;
 RA Milstein C., Frangione B.;
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2";
 RL Biochem. J. 121:217-225(1971).
 RN [10]
 RP DISULFIDE BONDS.
 RX MEDLINE=69064124; PubMed=5782707;
 RA Frangione B., Milstein C., Pink J.R.L.;
 RT "Structural studies of immunoglobulin G";
 RL Nature 221:145-148(1969).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J00230; AAB59393.1; -;
 DR PIR; A02148; G2HU.
 DR HSSP; P01857; 1FC1.
 DR GENE; HGNC:5526; IOHG2.
 DR MIN; 147110; -;
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00410; Ig_like; 1.
 DR SMART; SM00407; IGc1; 2.
 DR PROSITE; PS00290; IG_MHC; 2.
 DR Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1
 FT DOMAIN 1 98
 FT CH1.
 FT HINGE.
 FT DOMAIN 99 110
 FT DOMAIN 111 219
 FT DOMAIN 220 326
 FT CH3.
 FT INTERCHAIN (WITH A LIGHT CHAIN).
 FT INTERCHAIN (WITH A HEAVY CHAIN).
 FT INTERCHAIN (WITH A HEAVY CHAIN).
 FT INTERCHAIN (WITH A HEAVY CHAIN).
 FT INTERCHAIN (WITH A HEAVY CHAIN).
 FT AT OR NEAR THE COMPLEMENT-BINDING SITE.
 FT REMOVED POST-TRANSLATIONALLY (PROBABLE).
 FT S -> A (IN MYELOMA PROTEINS TIL & ZIE).
 FT /FTID=VAR_003889.
 FT C -> S (IN REF. 3).
 FT CONFLICT 109 109
 FT SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;
 SQ
 Query Match 10.7%; Score 378; DB 1; Length 326;
 Best Local Similarity 22.2%; Pred. No. 1.6e-16;
 Matches 123; Conservative 56; Mismatches 113; Indels 262; Gaps 14;
 QY 121 AKTTPSVYPLAPGSAATNSVTLGCLVKGVEPPTVTWNSGSLSSGVHTFPVAVLOSD 180
 DB 1 ASTKGPSVFPPLACSRSTSESTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLOSS 60
 QY 181 -LYTSSSVTPSPSTWPTVCNVAHPASSTKVDKIVPR-----DSGGPSEK 228
 DB 61 GLYSLSVTVTPSSNFGQTYTCNVNDRKPSNKTVDKTVKRCCEPCPCAPPVAGPS-- 118
 QY 229 SEINEKDLRKSLQGTALGNLKOIYYNSKATTSSEKSDAQFTLNTLLPKGFPTGHPW 288
 : : : : :
 : : : : :

Db 119 -----VFLFPPKP-----KDTLMI----- 132
 QY 289 YNDLLVLDLSTAAATSEYEGSSVDLYGAYGYQCAGGTENKTACMYGGVTLHDNNRLTEEK 348
 Db 133 -----SRTFVT-CVVVDVS-----HEDP 150
 QY 349 KVPINLWIDGKQTTVPIDKVKTSKEVTQVQLDQARHYLHGKFGLYNSDSFGGKVQRL 408
 Db 151 EVQFNWYVDG---VEVHNATKPRE----- 172
 QY 409 IVFHSSEGSTVSYDLFDAQGQYPTDLLRIYRDNTTISSTLSISLYLYTTSIVMTQTPTS 468
 Db 173 -----EQFNSTFRVSVLTVVHOD----- 191
 QY 469 LLVSAGDRVTITCRASQSVSDNVAVYQOKPGSKLLISYTSRVRAGVYDPDFSGSGSYGTD 528
 Db 192 -----WLNKEYKC-----KVSNNKGLPAP 210
 QY 529 FTLTSSVQAEDAAYVFCQDYNSPPTFGGKLEIKRADAAPTYSIFPPSSEQLTSGGA 588
 Db 211 IKTISKTKGO-----PREQVYTLPPSREEMTKNOV 242
 QY 589 SVVCFNNFYKIDNVKWKIDGSEKQ-----GVLSWTQDQSKDSTYSMSSTLTLTDE 643
 Db 243 SLTCLVKGFPDIAVEWESNGQPNENYKTTTPMLDS-----DGSFELYSKLTVDKSR 295
 QY 644 YERHNSYTCATHK 657
 Db 296 WQGNVFCSCVWHE 309
 RESULT 49
 KVLH_HUMAN
 ID KVLH_HUMAN STANDARD; PRT; 108 AA.
 AC P01605;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-1 region Lay.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=77038198; PubMed=824717;
 RA Capra J.D., Klapper D.G.;
 RT "Complete amino acid sequence of the variable domains of two human
 RT IgM anti-gamma globulins (Lay/pom) with shared idiotypic
 RT specificities";
 RL Scand. J. Immunol. 5:677-684(1976).
 CC -I- MISCELLANEOUS: THE SECOND AND THIRD HYPERVARIABLE REGIONS OF THIS
 CC CHAIN ARE IDENTICAL WITH THOSE OF THE HUMAN POM V-III KAPPA CHAIN,
 CC WITH WHICH IT SHARES CERTAIN IDIOTYPIC DETERMINANTS.
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
 CC GLOBULIN ACTIVITY.
 DR PIR; A01871; KLHULY.
 DR HSSP; P01607; 1REI.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGv; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 23
 FT DOMAIN 24 34
 FT COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 35 49
 FT COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 50 56
 FT COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 57 88
 FT COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 89 97
 FT COMPLEMENTARITY-DETERMINING-4.
 FT DOMAIN 98 107
 FT DISULFID 23 88
 FT NON_TER 108 108
 FT SEQUENCE 108 AA; 11834 MW; 739993A95431434A CRC64;
 SQ

THIS PAGE BLANK (USPTO)


```
Db 20 QVLOQSGAELARPAGSARVLSCKASGYTFTGYGVSWKQRTQGGLEWGEIYYPGSGNTYY 79
Qy 61 NOKFYDKATLTVDKSGSTAYMELRLSLTSDSAVYICARSTMTITNVMDYWGQGTSTVYSS 120
Db 80 SEKFKGKATLTITDKSSSTAYMHLSELTSDESAVYICARSSYYSDLFAYWGQGTTLTVSA 139
Qy 121 AKTTPSPVYPLAPGSAQAQNSMVTGLGKVGKGYFPEPVTVTWNSGSLSSGVHTFPFAVLQSD 180
Db 140 AKTTPSPVYPLAPGSAQAQNSMVTGLGKVGKGYFPEPVTVTWNSGSLSSGVHTFPFAVLQSD 199
Qy 181 LYTLLSSVTVPSTWSPSETVTCNVAHPASSTKVDKKIYPRDSGGPSEKSEINEKDLRKK 240
Db 200 LYTLLSSVTVPSTWSPSETVTCNVAHPASSTKVDKKIYPRDSGGPSEKSEINEKDLRKK 242
Qy 241 SELOQTALGNLKOIYYNSKAITSEKSAQDQFLTNLLFKGFTGHPMYNDLLVDLGSTA 300
Db 243 CRPCICTVPEVSSVFIFFPK-----PKDVLITL----- 271
Qy 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLLHDNNRLTEKKVPINLWIDGKQ 360
Db 272 -----TP-KVTCVVVDIS-----KDDPEVQSFWEVD--- 296
Qy 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVQRGLIVF 411
Db 297 -DVEYHTAQTPREPQFNSTFRSVSELPIHMDWDLNGREFKCRVNSAAPPAPIER----- 350
Qy 412 HSEGSTVSYDLFDAQGOVDPDLLLLRIYRDNNTTISSTLSLSISLYLTTISVMTQTPSLIV 471
Db 351 -----TISKT----- 355
Qy 472 SAGDRVTITCKASQSVNDVANYQKQPGSKLLISYTSRRYAGVPDRFSGSGYGTDFTL 531
Db 356 ----- 355
Qy 532 TISSVQAEDAAYVFCQDYNSPPTGGGKLEIKRADAAPTVIFPPSSEQLTSGGASVV 591
Db 356 -----KGRKAPQVYITPPKPEQMAKDVKSLT 382
Qy 592 CFLNFPYKPDINVKKIDGSRQNGVLNSWTDDQSKDSTYSMSSTLTLTCKDEYERHNSVT 651
Db 383 CMITDFEFDITVEQWNGQPAEN-YKNTQPMIDT-DGSFYIYSLNVOKSWEAGNTFT 440
Qy 652 CEATHK 657
Db 441 CSVLHE 446

RESULT 2
Q931M4 ID Q931M4 PRELIMINARY; PRT; 260 AA.
AC Q931M4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enterotoxin P.
GN SEP OR SAV1948.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RZ SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiranatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RL aureus.;"
RL Lancet 357:1225-1240(2001).
```

```
DR EMBL; AP003364; BAB58110.1; -.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_strp_toxin; 1.
DR Pfam; PF02876; Strep_strp_toxin; 1.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 260 AA; 30016 MW; 15C2D36270FA8241 CRC64;

Query Match 26.9%; Score 948; DB 16; Length 260;
Best Local Similarity 76.4%; Pred. No. 2.2e-51;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 226 SEKSEINEKDLRKSELQGTALGNLKOIYYNSKAITSEKSAQDQFLTNLLFKGFTG 285
Db 28 SEKSEINEKDLRKSELQGTALGNLKOIYYNEAKTENKESHQDFLOHTILFKGFTD 87
Qy 286 HPWYNDDLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLLHDNNRLT 345
Db 88 HSWYNDDLVDLFDKSDIVDKYKGVKVDLYGAYGYQCAGGTPNKTACMYGGVTLLHDNNRLT 147
Qy 346 EEKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQ 405
Db 148 EEKVPINLWIDGKQNTVPLETVTNKKNVTVQELDLQARRYLQEKYLNLYNSDVFQGVQ 207
Qy 406 RGLIVFHSSEGSTVSYDLFDAQGOVDPDLLLLRIYRDNNTTISSTLSLSISLYLVT 458
Db 208 RGLIVFHSSTEPSVNDLFGAQQVSNLTLRIYRDNKTINSENMHIDIYLYTS 260

RESULT 3
Q8VC16 ID Q8VC16 PRELIMINARY; PRT; 238 AA.
AC Q8VC16;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 26.2 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RZ SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019760; AAH19760.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 238 AA; 26224 MW; 35EC08E3DE5414AD CRC64;

Query Match 26.4%; Score 929.5; DB 11; Length 238;
Best Local Similarity 81.8%; Pred. No. 2.7e-50;
Matches 180; Conservative 15; Mismatches 20; Indels 5; Gaps 2;

Qy 457 TTSIVMTQTPSTSLVSGADRVITITCKASQSV--SND---VAMVQKPGQSPKLLISYSS 511
Db 18 SSDVYMTQTPSLVPSLGDQASISCRSSQSLVHNSGNTYLHWYLOKPGQSPKLLIYKSN 77
Qy 512 RYAGVPDRFSGSGYGTDFTLTISVQAEDAAYVFCQDYNSPPTGGTCKLEIKRADAA 571
Db 78 RFGVPDRFSGSGGTDFTLKISRVEADLGVYFCSQSTHVPPTPGGTCKLEIKRADAA 137
Qy 572 TVSIFPPSSEQLTSGGASVVCFLNNFYKPDINVKWKIDGSRQNGVLNSWTDDQSKDSTY 631
```

```

Db 138 TVSIPTSPSSQLTSGASVYVCFNNFYPRDINVKWKIDGSRQNGVLSNWTQDQSKDSTY 197
QY 632 SMSSTLTTLTKDEYERHNSVTCETHKTSPTSPVKSFNRE 671
Db 198 SMSSTLTTLTKDEYERHNSVTCETHKTSPTSPVKSFNRE 237

RESULT 4
Q9D8L4
ID Q9D8L4 PRELIMINARY; PRT; 473 AA.
AC Q9D8L4
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 181006009Rik protein.
GN IGH-1 OR 181006009Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Araiwa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK007918; BAB25349.1; -.
DR HSSP: P01842; 7FAB.
DR MGD: MGI:96443; Igh-1.
DR InterPro: IPR003599; Igh.
DR InterPro: IPR003597; Igh-cl.
DR InterPro: IPR003600; Igh-like.
DR InterPro: IPR003006; Igh.MHC.
DR InterPro: IPR003596; Igh_Y.
DR Pfam: PF00047; Igh_4.
DR SMART: SM00409; Igh; 2.
DR SMART: SM00407; Igc1; 3.
DR SMART: SM00406; Igv; 1.
DR SMART: SM00410; Igc-like; 1.
DR PROSITE: PS00290; IGH.MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FB8 CRC64;

Query Match 26.2%; Score 924.5; DB 11; Length 473;
Best Local Similarity 33.8%; Pred. No. 1.4e-49;
Matches 239; Conservative 53; Mismatches 123; Indels 293; Gaps 15;

QY 1 EVQLQSGDPLVKPGASVKISCKASGYSTGYIMHWKSPGKGLWIGRIINPNNGVTLY 60
Db 20 QVQLKSGAEIYKPGASVKISCKASGYSTGYIMHWKSPGKGLWIGRIINPNNGVTLY 79
QY 61 NQKFDKATLTVDKSTTAYMELRSITSDSAVYICARSTMTITNYMD---YWGQGSTVY 117
Db 80 NEKFKGATLTADKSSSTAYMQLSLTSDSAVYFCARS-----GYDYDFWYWGQGTLYT 135

```

```

QY 118 VSSAKTTTSPVYPLAPGSAQAQNSMTVLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVL 177
Db 136 VSAAKTTAPSVPYPLAPGCGTGTSSVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVL 195
QY 178 QSDLYTLSSVTVPSSTWPSSETVTCNVAHPASSTKVDKIVPRDGGSGGSEKSEENKRD 237
Db 196 QSGLYTLSSSVTVTNTWPSQTITCNVAHPASSTKVDKIEPR----- 238
QY 238 RKKSELQGTALGNLKOIYYVNSKAITSSSEKADQFLNTLLFKGFFTHGHPWYNLLVDLG 297
Db 239 ----- 238
QY 298 STAATSEYEGSVLDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEBEKVPINLWID 357
Db 239 -----VPI----- 241
QY 358 GKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVGRLIVFHSSEGS 417
Db 242 ---TQNPCPLKECPCAA-----PDLGGP----- 264
QY 418 TVSYDLFDAQGOVPTDLLRIYRDNTTISSTLSLSISLYLVTTSIVMTQTPTSLVSAGDRV 477
Db 265 ----SVFIFPPKIKDVL-----ISLSPWTCVV----- 290
QY 478 TITCKASQSVSDVAVYQOQPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTSSVQ 537
Db 291 -----DPDVOISW-----FVNNVE 309
QY 538 AEDAAVYFCQDYNSPPTFGGTTKLEIKRAD----- 568
Db 310 VHTAQTQTHREDYNS--TLRVVSALPIQHDWMSGKEFKCKVNNRALPSPIEKTSKPRG 367
QY 569 --AAPTSTVSIFFSSPQLTSGGASVVCFLNFKYPKDINVKWKIDGSRQNGVLSNWTQDQ 626
Db 368 PYRAQVYVLPPEAEMTKKESLTCMGTGFLPAELAVDWTNGRTEQN-YKNTATVLDS 426
QY 627 KSTYSMSSTLTTLTKDEYERHNSVTCETHK-----TSTSPVKSFN 669
Db 427 -DGSYFMYSKLRVQKSWERGSFLACSVVHEGLHNLTKTISRSLGK 473

RESULT 5
Q99SU3
ID Q99SU3 PRELIMINARY; PRT; 260 AA.
AC Q99SU3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Enterotoxin P.
DE SEP OR SAI761.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003135; BAB43036.1; -.
DR HSSP: P13163; 1SYT.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam: PF01123; Staph_Strep_toxin; 1.
DR Pfam: PF02876; Staph_Strep_tox_C; 1.
DR PRINTS: PR00279; BACTRILTTOXIN.

```

DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Complete proteome.
SQ SEQUENCE 260 AA; 29708 MW; 087C5B4EC028CFDB CRC64;

Query Match 26.1%; Score 918; DB 16; Length 260;
Best Local Similarity 73.0%; Pred. No. 1.6e-49;
Matches 170; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKSELOGTALGNLQKIYYIYNKAITSSSEKSAOFLTNLLFKGFTG 285
DB 28 SEKSEINEKDLQKSELOGTALSRLQTYIHNGSAITENKESNDQFLKNTILFNDFFTG 87
QY 286 HPWYNLLVDLSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 345
DB 88 HQWYNLLVDLSDKDTANIYKGGKVDLYGYGYOCTGGTFFKTACMYGGVTLHDNNQLE 147
QY 346 EEKKVPINLWIDGKQTPIDPKVYKSKKEVTVOELDLQARHYLHGKFLYNSDSFGKQV 405
DB 148 EEKKVPINLWIDGKQTPVPLGTVKTKKKEVTVOELDLQSRHYLHETYNLYNDAENGKIQ 207
QY 406 RGLIVPHSSEGSTVSVDLFDAGQYPTDLLRIYRONTTISSTLSISLYLYTT 458
DB 208 RGLIEFHPSSGDSVGYDLFGAQQYPTQLRIYRONTTKISKNMHIDIYLYTT 260

RESULT 6
Q99L25 PRELIMINARY; PRT; 473 AA.
AC Q99L25;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 181060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003888; AAH03888.1; -.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; Igc1; 3.
DR SMART; SM00406; Igc1; 1.
DR SMART; SM00410; Ig-like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 52449 MW; BE9889B79860A155 CRC64;

Query Match 26.1%; Score 917.5; DB 11; Length 473;
Best Local Similarity 33.8%; Pred. No. 3.8e-49;
Matches 232; Conservative 59; Mismatches 141; Indels 255; Gaps 12;

QY 1 EVOLOQSGDLVKPGASVKISKACSGYFTGYMHWKQSPGKGLWIGRINPNNGVTLY 60
DB 20 QVLOQSDAELVKPGASVKISKVSGYFTDTHIHWKQRPQGLEWIGYIPRGSTKY 79
QY 61 NQFKDKATLVKDSSTTAYMELRSLTSEDSAVYYCARSTMITN---YVMDYWGQGTSV 116
DB 80 NEKFGKATLTADKSSATYMLNSLTSEDSAVCFCSRGSIIYGYGLYFDYWGQGTI 139
QY 117 TVSSAKTTPPSVYPLAPGSAAGTNSWVTLGCLVKGFFPEPVTVWNSGSLSGVHTFPFV 176
DB 140 TVSSAKTTPPSVYPLAPVCGDGTGTSVTLGCLVKGFFPEPVTVWNSGSLSGVHTFPFV 199

QY 177 LOSDLTYLSSSVTPSSTWSPSETVCNVAHPASSTKVDKKIVPRDSGGPSEKSEBINEKD 236
DB 200 LOSDLTYLSSSVTPSSTWSPSITCNVAHPASSTKVDKKIEPR---GPTIKPCP----- 251
QY 237 LRKKELOGTALGNLQKIYYIYNKAITSSSEKSAOFLTNLLFKGFTGHPWYNLLVDL 296
DB 252 ----- 251
QY 297 GSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWI 356
DB 252 -----PCKCPAPN-----LLGGPSVF----- 267
QY 357 DGKQTTPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFGKQVORGLIVPHSSEG 416
DB 268 -----IFPPKIKDVLMI---SLS 282
QY 417 STVSVDLFDAGQYPTDLLRIYRONTTISSTLSISLYLYTTISYMTOTPTSLLSVAGDR 476
DB 283 PMVTCVWVDVSEDDPD-----VQISWFVNNVEVLTAQTQT----- 317
QY 477 VTITCKASQSVSNDAVYQOKPGQSPKLLISYTSRYAGVPRDFSGSGYGTDTLTITISSV 536
DB 318 -----HREDYNSTLRVV 329
QY 537 QA-----EAAVYFCQDYNSPTFTGGTKLEIKRADAAPTVSIPPPSSEQLTSGA 588
DB 330 SALPIQHODWMSGKFEKCKVNNKALPAPIRTISKPKSVRAPOVYVLPPEEMTKKQV 389
QY 589 SVVCFLNFPKIDINVWKIDGSERON-----GYLNSWTDQDSKDSYSTSMSSTLTLTDE 643
DB 390 TILTCMVTDFMPEDIYVEWTNNGKTELNYKNTPEVLDS-----DCSYFMYSKLRVEKN 442
QY 644 YERHNSYTCETHK--TSTSPIVKSEN 669
DB 443 WVERNSYSCVVVHEGLHNHHTKSESR 469

RESULT 7
Q8VC55 PRELIMINARY; PRT; 239 AA.
ID Q8VC55;
AC Q8VC55;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hypothetical 26.3 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021781; AAH21781.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; Igc1; 1.
DR SMART; SM00406; Igc1; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26303 MW; C16119CACA25C337 CRC64;

Query Match 25.8%; Score 910; DB 11; Length 239;
Best Local Similarity 76.0%; Pred. No. 4.4e-49;
Matches 177; Conservative 21; Mismatches 25; Indels 10; Gaps 2;

QY 449 LISISLYLT-----TSIVMTQTPTSLLSVAGDRVTITCKASQSVSN-----VAWYQOKP 498
DB 449 LISISLYLT-----TSIVMTQTPTSLLSVAGDRVTITCKASQSVSN-----VAWYQOKP 498

DR	PROSTE; PS00290; IG_MHC; UNKNOWN_1.
KW	Hypothetical protein.
FT	NON_TER 1
SQ	SEQUENCE 233 AA; 25781 MW; BLCI84DAL49A16EB CRC64;
Query Match 25.6%; Score 900; DB 11; Length 233; Best Local Similarity 81.6%; Pred. No. 1.8e-48; Matches 173; Conservative 14; Mismatches 25; Indels 0; Gaps	
QY	I VMTQTPTSLLVAGSDRVTTTCASQSVDVAWYOQKPGSQPKLLISYTSSRYAGVPSR 519
DB	: : : : : : :
DB	21 IQMTQTSSLVASLGDVRTTISCSSQGIANYLNWYOQKPDGCTVKLLIYTSSLSHGVPSR 80
QY	F SSGGYGTDTLTLISSVQAEDAAVFQQDYNSPPTFGGGTKLEIKRADAPTVSIFPPS 520
DB	: : : : : : :
DB	81 FSGSGGTDYSLTISNLEPEDIAITYYCQYRLPLPWTFGGGKTLEIKRADAPTVSIFPPS 140
QY	S EQLTSGGASVCFNLNFYPKPDINVKWKIDGSRONGVLNSWTDDQDSKDSTYSMSSTLT 639
DB	: : : : : : :
DB	141 SEQLTSGGASVCFNLNFYPKPDINVKWKIDGSRONGVLNSWTDDQDSKDSTYSMSSTLT 200
QY	T KDEYERHNSYTCETHAKTSTSPIVKSFNRE 671
DB	: : : : : : :
DB	201 TKDEYERHNSITCEATHKTSTSPIVKSFNRE 232
RESULT 12	
OR062	PRELIMINARY; PRT; 234 AA.
ID	QR062
AC	QR062;
DT	01-JUN-2002 (TrEMBLrel. 21, Created)
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Hypothetical 25.9 kDa protein.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	RNCB_TaxID=10090;
NB	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=COLON;
RA	Straussberg R.;
RL	Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases .
DR	EMBL; BC027418; AAH27418.1 ;
KW	Hypothetical protein.
SQ	SEQUENCE 234 AA; 25857 MW; 4EB08C81426AEAB1 CRC64;
Query Match 25.4%; Score 896; DB 11; Length 234; Best Local Similarity 81.6%; Pred. No. 3.2e-48; Matches 173; Conservative 13; Mismatches 26; Indels 0; Gaps	

[illegible]

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:18977).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC012207; AAH12207.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_2.
SQ SEQUENCE 278 AA; 29778 MW; F894F955DDCD948A CRC64;

Query Match 25.4%; Score 893; DB 11; Length 278;
Best Local Similarity 76.8%; Pred. No. 6.1e-48;
Matches 169; Conservative 18; Mismatches 33; Indels 0; Gaps 0;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYMHVWKSPGKGLGWIGRINPNNGVTLY 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 QVLOQPGALVKPGASVKLSCKASGYFTSYMHVWKRRGQGLGWIGNINPNSGNTY 79
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NOKFKDKATLVVDKSTTAYMELRSLTSEDSAVYYCARSTMITNYYMDYWGQGTSTVSS 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 NEXFKNKATLAVDKSSSTVMQLSSTSEDSAVYYCTRGVYDDVYFVWGAGTTVTSS 139
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 AKTTPSVVPLAPGSAQAQNSMTVLCGLVKGYPEPVTYTNWNSGSLSSGVHTFPVQLSD 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
140 AKTTPSVVPLAPGCGGTTGSSVTLCGLVKGYPEPVTYTNWNSGSLSSGVHTFPALLQSG 199
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 181 LYTLLSSSVTPSPSETVTCNVAHPASSTKVDDKKIVPR 220
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 LYTLLSSSVTPSPSETVTCNVAHPASSTKVDDKKIEPR 239
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024405; AAH24405.1; -.
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match 25.2%; Score 889; DB 11; Length 469;
Best Local Similarity 32.9%; Pred. No. 2.2e-47;
Matches 221; Conservative 68; Mismatches 129; Indels 254; Gaps 16;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYMHVWKSPGKGLGWIGRI-NPNNGVTLY 59
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 EVNLVESGGVLQPGGSLRLSCAASGFTFDYYSNVRPPGKALEWLGFIKRNKANGYT 79
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 60 -YNQFKDKATLVVDKSTTAYMELRSLTSEDSAVYYCARSTMITNYY----MDYWGQGT 114
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 EYSASVKGRETTISRDNSQSILYQMLNALRAEDSATIYCARRDRSSYYISGTSFAYWGQT 139
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 115 SVTVSSAKTTPSVVPLAPGSAQAQNSMTVLCGLVKGYPEPVTYTNWNSGSLSSGVHTFP 174
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
140 LVTVAATKTPSVVPLAPGSAQAQNSMTVLCGLVKGYPEPVTYTNWNSGSLSSGVHTFP 199
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
QY 175 AVLQSDLYTLSSSVTPSPSETVTCNVAHPASSTKVDDKKIVPRDSGGSPSEKSEINE 234
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 AVLQSDLYTLSSSVTPSPSETVTCNVAHPASSTKVDDKKIVPRDCG----- 248
QY 235 KDLRKSELOQTALGNLKOIYYNSKAITSSSEKSAQDQFLTNLLFKGFTTGHWPYNDLLV 294
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 -----CKPCTCTPEVSSVFIPPK-----PKDVLITL----- 277
QY 295 DLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMGYVTLHDNNRLTEKKVPINL 354
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
278 -----TP-KVTCVVVDIS-----KDDPEVQFSW 299
QY 355 WIDGKQTTVPIDKVTSSKEV-----TVQELDLQARHYLHGK-EGL-YNDSDFGKQVQ 405
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 FVD---DVEVHTAQTKPREQFNSTERSVSELPIMHQDWLNGKEFKRVNSAAPPAPIE 355
QY 406 RGLVPHSSEGSSTVSYDLFAQGOYPTDLLRIYRDNTTISSTLSLSLYLYTTSIVMTQT 465
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 K-----TISKT----- 361
QY 466 PTSLLYSAGDRVITITCKASQSVNDVAVYQKPGQSPKLLISVTSRYAGVPDRFSGSY 525
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
362 -----KGRPKAPQVYTIPTPPKEQMAK 382
QY 586 GGASVVCFLNFPKIDNVKWKIDGSEKONGVLSNWTDDSKDSTYSMSSTLTLTDEYE 645
DB :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
383 DKVSLTCTMTDFEPDITVEWQWNGQPAEN-YKNTQPIQMDT-DGSYFYVSKLVNQKSNWE 440
QY 646 RHNSYTCETHAK 657
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
441 AGNTFTCSVLHE 452

RESULT 15
Q8R3H6 PRELIMINARY; PRT; 474 AA.
AC Q8R3H6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 51.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025447; AAH25447.1; -.
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 51748 MW; 8608B57C6CD2874A CRC64;

Query Match 25.2%; Score 888.5; DB 11; Length 474;
Best Local Similarity 35.0%; Pred. No. 2.4e-47;
Matches 234; Conservative 69; Mismatches 124; Indels 241; Gaps 18;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYMHVWKSPGKGLGWIGRINPNNGVTLY 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 QVOLLQSGDELVKPGASVKISCRASGYAFSKSWNVVWRPGRPGKLEWIGRIFFGDGDTHY 79
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 NOKFKDKATLVVDKSTTAYMELRSLTSEDSAVYYCARSTMITNYYMDYWGQGTSTVSS 120
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 SGKFOGKAKLTADKSSVTAFQLTSLTSEDSAVYFCAKSDYGDY-FDDWGQCATVTSS 138
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 AKTTPSVVPLAPGSAQAQNSMTVLCGLVKGYPEPVTYTNWNSGSLSSGVHTFPVQLSD 180
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 AKTTPSVVPLAPGCGGTTGSSVTLCGLVKGYPEPVTYTNWNSGSLSSGVHTFPALLQSG 198
```



```
Query Match          24.1%; Score 850; DB 11; Length 234;
Best Local Similarity 76.9%; Pred. No. 2.3e-45;
Matches 163; Conservative 17; Mismatches 32; Indels 0; Gaps 0;

QY 460 IYMTQPTSLLSAGDRVTITCKASQSVNDVAVYQKPCQPKLLISYTSRYAGVPPR 519
DB 22 IQLTQSPASLSASVGETVITCRASENIYSYLAWYQKQKSPQLLVYNAKTLADGVPSR 81
QY 520 FSGSGYGTDTLTLSVQAEADAAVFCQDYNPPTFGGKTKLEIKRADAAPTYSIFPPS 579
DB 82 FSGSRGTQFSLKINSLOPEFGSYCQHSGIPFTFGSGTKLEIKRADAAPTYSIFPPS 141
QY 580 SEQLTSGGASVVCFLNFPKIDNVKWKIDGSRQNGVLNSWTDQDSDKSTYSMSSTLTL 639
DB 142 SEQLTSGGASVVCFLNFPKIDNVKWKIDGSRQNGVLNSWTDQDSDKSTYSMSSTLTL 201
QY 640 TKDEYERHNSYTCEATHKSTSTSPIVKSFNRE 671
DB 202 TKDEYERHNSYTCEATHKSTSTSPIVKSFNRE 233

RESULT 21
Q91205 PRELIMINARY; PRT; 473 AA.
ID Q91205;
AC Q91205;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 51.9 kDa protein.
GN AU044919.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010327; AAH10327.1;
DR MGD; MGI:2144967; AU044919.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

Query Match          22.8%; Score 803; DB 11; Length 473;
Best Local Similarity 32.0%; Pred. No. 4.9e-42;
Matches 214; Conservative 74; Mismatches 138; Indels 242; Gaps 18;

QY 1 EVLOQSGDPLVKPGASVKISKASGYSTGYMHVVKVSPKGLWIGRIINPNNGVTLY 60
DB 20 EVLVESGGLVKPGGSRKLSAAGSTEDYCMHVRQAPKEGLWAYVINSGSTTYY 79
QY 61 NQKFKDKATLTVDKSSTTAYMELRSLTSDSAVYYCARSTMTINVMYDYGQTSVTSS 120
DB 80 ADTVKGRFTISRDNANKTLFQLMRSURSDTAMYYCARELWLR--IDYWGQGTITVSS 137
QY 121 AKTTPSVYPLAPGSAQTNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPVAVLQSD 180
DB 138 AKTTPSVYPLAPGCGDTTGSSTVTLGLCLVKGYFPEPVTVTWNSGSLSSSVHFPALQSG 197
QY 181 LYTSSSVTPSSTWPSSETVTCNVAHPASTVKYDKKIVPRDSGGPSEKSEINE-----KD 236
DB 198 LYTMSSTVTPSSTWPSSTVTCNVAHPASTVTDKKLEP---SGP---ISTINPCPCKE 251
QY 237 LRK--KSEQLQGTALGNLKOIYYNSKAITSSSEKSADQFLNTLLFGFTGHPWYNDLLV 294
DB 252 CHKCAPNLEG---GPSVFIPFPNIKDV-----LMI 279
```

```
QY 295 DLGSTAATSEYEGSSVDLYGAYGYQACGTPNKTACMYGGVTLHDNRLTEKKVPINL 354
DB 280 SL-----TP----- 283
QY 355 WIDGKQTTVPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRLIVFHSS 414
DB 284 ---KVTQVVD---VSEDDPDVQISWVFN--VEVHTA 313
QY 415 EGSTVSVDLFDAGQYPTDLLRIYRDNNTTISLSLSLYLTTSIVMTQTPTSLLSVAG 474
DB 314 QQTTHREDY-----NSTIR-----VYSA- 331
QY 475 DRVTITCKASQSVNDVAVYQKPGSKLLISYTSRVAGVDPDRFSGCYGCTDFTLTIS 534
DB 332 -----LPIQHDMWSKEFKCKYN 350
QY 535 SYQAEAAVYFCQDYNPPTFGGKTKLEIKRADAAPTYSIFPPSSEQLTSGGASVVCFL 594
DB 351 N-----KDLPSPIE---RTISKIKGLVRAPQVYILLPPPAEQLSRKDVSUTCLV 395
QY 595 NNFPKIDINVKWKIDGSRQN-----GVLSNWTQDQSDKSTYSMSSTLTLTKDEYERHNS 649
DB 396 VGFNPGDISVENTSNGHTEENYKDTAPVLDS-----DGSYFIYSKLDIKTSKNEKTD 448
QY 650 YTCSEATHK 657
DB 449 FSCNVRHE 456

RESULT 22
O85217 PRELIMINARY; PRT; 268 AA.
ID O85217;
AC O85217;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Enterotoxin J.
GN SEJ.
OS Staphylococcus aureus.
OG Plasmid pIB485.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KS11410;
RX MEDLINE=99052098; PubMed=9835033;
RA Zhang S., Iandolo J.J., Stewart G.C.;
RT "The enterotoxin D plasmid of Staphylococcus aureus encodes a second
RT enterotoxin determinant (sej).";
RL FEMS Microbiol. Lett. 168:227-233(1998).
DR EMBL; AF053140; AAC78590.1;
DR HSSP; P13163; IESF.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; BACTRLTOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Plasmid.
SQ SEQUENCE 268 AA; 31230 MW; AC5F3546060ACE22 CRC64;

Query Match          21.8%; Score 768; DB 2; Length 268;
Best Local Similarity 58.3%; Pred. No. 3.4e-40;
Matches 141; Conservative 39; Mismatches 62; Indels 0; Gaps 0;

QY 228 KSEINEKDLRKSEKLGQGTALGNLKOIYYNSKAITSSSEKSADQFLNTLLFKGFTGHP 287
DB 27 KNETIKEKNLHKSSLSITLNLRIHYFFNEKGISEKIMTEQFLDYTLLEKSFISHS 86
QY 288 WYNDLLVLGSLTAATSEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNRLTEE 347
DB 87 QYNDLLVQFSDSKETNFKGKQVDLYGSYGYQCSCGKPKNTACMYGGVTLHENNQLYDT 146
```


Db 346 -----KCKVSNK-----GLP-- 355
QY 520 FSGSGYGTDTLTITSSVQAEADAAVFCQDYNPPTFGGKLEIKRADAAPTVSIFPPS 579
Db 356 -----SSIEK-----TISKAGQPREPOVYTLPPS 380
QY 580 SEQLTSGGASVVCFLNFFPKDINVKWKIDGSRQN-----GVLNSWTQDQSKDSTYSMS 634
Db 381 QEMTKNQVSLCLVKYGFPSDIAEVESNGQPNKYKTPPVLDSE-----DGSFFLY 433
QY 635 STLTTLTKDEYERHNSYTCBATHK 657
Db 434 SRLTVDKSRWQGNVFCSCVMHE 456

RESULT 25

Q8TCD0
ID Q8TCD0 PRELIMINARY; PRT; 239 AA.
AC Q8TCD0;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hypothetical 26.2 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC022362; RAH22362.1; -
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26234 MW; FAGEDC3A3B03871D CRC64;

Query Match 18.6%; Score 654.5; DB 4; Length 239;
Best Local Similarity 58.1%; Pred. No. 3.3e-33;
Matches 126; Conservative 34; Mismatches 52; Indels 5; Gaps 1;
QY 460 IVMTQPTSLVSGADRVITCKASQV-----SNDVAVQKQKPGSPKLLISYSSRYA 514
Db 22 VVMTQSLPLVTLGQPASISCRSTQSLVSDGNTVLNWFQORPGQSPRLIYKVSNRDS 81
QY 515 GVPDRFSGSGYGTDTLTITSSVQAEADAAVFCQDYNPPTFGGKLEIKRADAAPTVS 574
Db 82 GVPDRFSGSGYGTDTLTITSSVQAEADAAVFCQDYNPPTFGGKLEIKRADAAPTVS 141
QY 575 IFPPSSEQLTSGGASVVCFLNFFPKDINVKWKIDGSRQNGVLNSWTQDQSKDSTYSMS 634
Db 142 IFPPSDEQLKSGTASVVCFLNFFPREAKVQMKVDNALQSGNSQESVTEQDSKDSYLS 201
QY 635 STLTTLTKDEYERHNSYTCBATHKSTSPVKSFNRE 671
Db 202 STLTLSKADYERHKVYACEVTHQGLSSPVTKSFNRGE 238

RESULT 26

Q8VCX7
ID Q8VCX7 PRELIMINARY; PRT; 613 AA.
AC Q8VCX7;
DT 01-NAR-2002 (TremBLrel. 20, Created)
DT 01-NAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hypothetical 67.9 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SALIVARY GLAND;

RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC018315; AAH18315.1; -
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; IgC1; 4.
DR SMART; SM00406; IgV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67855 MW; 41A9384DD4C22862 CRC64;

Query Match 17.6%; Score 618.5; DB 11; Length 613;

Best Local Similarity 28.1%; Pred. No. 2e-30;
Matches 196; Conservative 90; Mismatches 211; Indels 201; Gaps 28;

QY 1 EVOLQSGPDLVKPGASVKISCKASGYFTGYMHVWKSPGKGLWIGRINPNNGVTLY 60
Db 20 QVOLQSGAELMKPGASVKISCKATGYFTSSWIEWKORPGHGLEIILPGSGSTNY 79
QY 61 NQKFKDKATLTVDKSTTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGQSTVTSVS 120
Db 80 NEKFKGKATFTADTSSNTAYMQLSLTSDSAVYYCAR--RLGRWYFDVWGAGTTVTSS 137
QY 121 AKTTPPSVYPLAP-GSAAQTNSMVLGCLVKGYFPPVPTVW---NSGSLSSGVHTFPV 176
Db 138 ESQSPFNVPFLVSCESPLSKLVAMGCLARDLPSTISFTWNYQNNTVIOGIRTFPTL 197
QY 177 LQSDLYTLSSVTV-PSS--TWPSETVTCNVAH-----PA---SSTKVDKKIYP 219
Db 198 RTGCKYLATSOVLLSPKSILEGSDVLCIKHYGKNRDLHVPPIPAVAEMPNVNVFVP 257
QY 220 RDS-GGSPSEKSEINEKDLRKSELOGTALGNLQIYYNYSKAITSEKSAQDLNTLL 278
Db 258 RDGFGSPAP-----RKSCL-----ICEATNFTPKPITVS-----286
QY 279 FKGFTHGWYND-LLVDLGSTA--ATSEYEGSSVDLYGAYGYCAGGTPNKTACMYGG 335
Db 287 -----WLKDGKLVESGFTDPVTIENKGSTPQTY-----315
QY 336 VTLHNNRLTEKKVPINLWIDGKQTPVDIKVTSKKEVTVOELDLQARHVLHGKFLY 395
Db 316 -----KVIST---LTISEID-----WL 329
QY 396 NSDSFGKVGORGLIVPHSSEGSTVSVDLFDAGQYPTDLLRIYRDNTTISSTLSISLYL 455
Db 330 NLNVYTCRVDRHGLTFLKNVSTCA-----ASPSTDI-----361
QY 456 YTTISIVMTQPTSLVSGADRVITCKASQVSND---VANVQKQKPGSPKLLISYSSR 512
Db 362 ----LTFITPPFADIFLSKANLTCVLSNLTATYETLNISWASQS-GEPLTKIKIMESH 416
QY 513 YAGVPDRFSGSGYGTDTLTITSSVQAEADAAVFCQDYNPPTFGGKLEIKR 566
Db 417 PNGT---FSAKG-----VASVCVEDNNRKEFVCTVTHRDLPSP-----QKFKISK 459
QY 567 ADAA-----PTVSIFPPSPSEQLT-SGGASVVCFLNFFPKDINVKWKIDGS--ERQNGVLN 619
Db 460 PNEVIKHPAPVYLLPPAREQLNRESATVTCVKGFSPADISVQWLQRGOLLPOEKYVTS 519
QY 620 SWTDQDSKDSYMSSTLTTLTKDEYERHNSYTCBATHK 657
Db 520 APMPEGAPGFYFTHTSILTVEENNSGETYTCVVGHE 557

RESULT 27

Q96GA6
ID Q96GA6 PRELIMINARY; PRT; 614 AA.
AC Q96GA6;
DT 01-DEC-2001 (TremBLrel. 19, Created)

QY 593 FLNFPKIDINVKWKIDGSRQNG-----VLNSWTDQDSDKSDTSMSTLTLTAKDEYERHN 648
Db 386 LVRAFNPKEVLVRW-LHGNEELSPESYLVEPLKEPEGATTYLVTSVLRVSAEINKQGD 444
QY 649 SYTCEATHK 657
Db 445 QYSCMVGHE 453
RESULT 29
Q91WT1 PRELIMINARY; PRT; 481 AA.
AC Q91WT1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical 52.1 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013490; AAH13490.1; -
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 481 AA; 52105 MW; 97DF68D159463F65 CRC64;

Query Match 16.6%; Score 583; DB 11; Length 481;
Best Local Similarity 26.9%; Pred. No. 2.3e-28;
Matches 179; Conservative 73; Mismatches 167; Indels 246; Gaps 21;

QY 1 EVQLQSGPDLVKPGASVKISKASGYFTGYIMHWKSPGKGLWIGRINPNNGVTLY 60
Db 20 QVQLQSGPELVKPGASVKISKASGYFTSYIHWKRPQGLWIGWIVPGDNTKY 79
QY 61 NQKFKDKATLTVDKSTTAYMELRSLTSDSAVYICARSTMTINYVMDYWGQTSVTSS 120
Db 80 NEKFQKATLTADKSSSTAYMELSLTSDSAVYICFTRG---GGWAFDYGQGTTLTVSS 136
QY 121 AKTTPSVVPLAPGSAQAQNSMVTLCGLVKGYFPE-PVTVTW-NSGSLSSGVHTFPVAVLQ 178
Db 137 EPAREPTIYPLTPQALSSDPVI-IGCLLHDYFPSTGMVMTWKGSKGDKITTVNFPALAS 195
QY 179 SDLYTLSSSVTPSPSTWP-SETVTCNVAHPASSTKVDKKIVPRDGGPSEKSEEEINEKDL 237
Db 196 GGRYTMSSQLTPAVCEPGESEKSCVQH-----DSNPVQELNV--- 234
QY 238 RKKSELOGTALGNLKOIYYNSKATITSSEKSAQDFLTNTLLFKGFTGHPWYNDLLVDLG 297
Db 235 -----NCPGIC---SPPTPPPPSCQPSLS-----LQPALEDLL--LG 268
QY 298 STAATSEYEGSSVDLYGAYGYQCAGGTNKTACMYGGVTLHDNNRLTEKKVPINLWID 357
Db 269 SDASIT-----CTLNG----- 279
QY 358 GKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVGQRLIVFHSSEGS 417
Db 280 ----- 279
QY 418 TVSYDLFDAQGYDPDLLRIYRDNMTTISLSLSISLYLTTSIVMTQTPTSLLSVAGDRV 477
Db 280 -----LRDPEG-----AVFTWEP-----TGKDAV 299
QY 478 TITCKASQSVNDVANYQKPGSPKLLISYTSRRYAGVDPDRFSGSGYGTDTLTLTSSVQ 537
Db 300 -----QKKAVQNSCGCY-----VSSLPLPGCAERWNS---GASFCKCTVTHPE 338

QY 538 ABDAVAVFQQQDYNPSPTFGGKLEIKRADAAPTVSIFPPSPSEQLTSGG-ASVVVCFLLN 596
Db 339 SDTLTGCTIAKVTNTFP-----PQVHLLPPPEELALNELVSLTCLVRA 382
QY 597 FYPKIDINVKWKIDGSRQNG-----VLNSWTDQDSDKSDTSMSTLTLTAKDEYERHNSYTC 652
Db 383 FNPKEVLVRW-LHGNEELSPESYLVEPLKEPEGATTYLVTSVLRVSAEINKQGOYSC 441
QY 653 EATHK 657
Db 442 MVGHE 446
RESULT 30
Q8VCV5 PRELIMINARY; PRT; 481 AA.
ID Q8VCV5;
AC Q8VCV5;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 52.3 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY TUMOR;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018455; AAH18455.1; -
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00409; Ig_3.
DR SMART; SM00407; Ig_c1; 3.
DR SMART; SM00406; Ig_v; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 481 AA; 52326 MW; 52B44C5826807143 CRC64;

Query Match 16.3%; Score 573; DB 11; Length 481;
Best Local Similarity 26.6%; Pred. No. 9.6e-28;
Matches 177; Conservative 83; Mismatches 159; Indels 246; Gaps 23;

QY 1 EVQLQSGPDLVKPGASVKISKASGYFTGYIMHWKSPGKGLWIGRINPNNGVTLY 60
Db 20 EIQLQSGPELVKPGTSVKVSKASGYFDYNIYWKSHGKSLWIGYIDPYNGSSY 79
QY 61 NQKFKDKATLTVDKSTTAYMELRSLTSDSAVYICARSTMTINYVMDYWGQTSVTSS 120
Db 80 NQKFKGATLTVDKSSNTAFMYLNLTSDSAFYICAREWYGAWEA--FWGGTLLTVSA 137
QY 121 AKTTPSVVPLAPGSAQAQNSMVTLCGLVKGYFPE-PVTVTW-NSGSLSSGVHTFPVAVLQ 178
Db 138 ESARNPTIYPLTPALSSDPVI-IGCLLHDYFPSTGMVMTWKGSKGDKITTVNFPALAS 196
QY 179 SDLYTLSSSVTPSPSTWP-SETVTCNVAHPASSTKVDKKIVPRDGGPSEKSEEEINEKDL 237
Db 197 GGRYTMSSQLTPAVCEPGESEKSCVQH-----DSNPVQELDV 235
QY 238 RKKSELOGTALGNLKOIYYNSKATITSSEKSAQDFLTNTLLFKGFTGHPWYNDLLVDLG 297
Db 236 ----- 235
QY 298 STAATSEYEGSSVDLYGAYGYQCAGGTNKTACMYGGVTLHDNNRLTEKKVPINLWID 357
Db 236 -----NCSGTP-----PPPI----- 246
QY 358 GKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVGQRLIVFHSSEGS 417
Db 358 ----- 417


```
Db 140 VTVSSGSASAPTLFPLVSCENSPSDTSS-VAVGCLAQDFLPDSITFSMKYKNNSSDISS-T 197
QY 171 HTFPAVLQSDLYTLSSVTVPPSS---TWPSSETVTCNVAHPASSTKVKKIIVPRDSCGPPSE 227
Db 198 RGFPVSLRGKYAATSOALLPKSKVMQGTDEHVCKVQHP-NGNKEKNVPLPVAIAELPK 256
QY 228 KSEENEKDLRKSELOGTALGNLQAIYYNSKAITSSSEKSADQFLTNTLFLKGFETGHP 287
Db 257 VSVFVPPRD-----QATGFSP-----GFF-GNP 271
QY 288 WYNLDLLVGLGTAATSEYEGSSVDLYGAYGYQCAGTTPNKTACMYGGVTLHDNNRLTEE 347
Db 272 RKSKLIC-----QATGFSP----- 285
QY 348 KVPINLNLWDGKQ--TTVPIDKVKTSKKE-----VTVOELDLQARHYLHGKFL 394
Db 286 ROIQVSWLRGKGVGGVTTDQVQAEAKESGPTTYKVTSTLTIKESD-----W 333
QY 395 YNSDSFGKGVQGLVIFHSSEGSTVSYDLFDAQGYDPTLLRIYRDNTTISTSLISLY 454
Db 334 LSQSFCTCRVDHRLGTLFQONASSMCVDP-----QDTAIRVF-----AIPPSFASIFL- 380
QY 455 LYTTTSIVMTQPTSLLSVAGDRVTITCKASQSVNDVAVYQKQPSKLLISYTSRYA 514
Db 381 --TKSTKLTLCLVTDL--TTYDSVII-----SWTRQN-GEAVK---THNINISES 420
QY 515 GVPDRFSSGVTGDTTLTISSVQAEADAAYFCQDYNSPPTF-----GGGKLEIK 565
Db 421 HPNATFSAVG-----EASI--CEDDWSNGERFTCTVTHTDLPSPKQTIS 463
QY 566 RADA--PTVSIFPPSSSEOLT-SGGASVVCFLNNFYPKDINVKKWIDGS--ERQNGVL 618
Db 464 RPKGVALLHRPDVYLLPAREQLNLRRESATITCLVTGSPADVFQVMQRGQPLSPERYVT 523
QY 619 NSWTDQDQSKDSTYSMSSTLTLTKDYEYRHNSYTCEATHK 657
Db 524 SAPMPEQAPGRYFAHSILTVSEEWNTGETYTCVVAHE 562

RESULT 33
Q9BQB8 PRELIMINARY; PRT; 597 AA.
AC Q9BQB8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Unknown (protein for MGC:1905) (protein for MGC:1228).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006180; AAH06180.1;
DR EMBL; BC001872; AAH01872.1;
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 4.
DR SMART; SM00406; IGV; 1.
```

```
DR SMART; SM00410; IG-like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
SQ SEQUENCE 597 AA; 65300 MW; 2DAFAD50A6375851 CRC64;

Query Match 15.5%; Score 545; DB 4; Length 597;
Best Local Similarity 25.9%; Pred. No. 7.le-26;
Matches 182; Conservative 103; Mismatches 213; Indels 204; Gaps 29;

QY 1 EVQLQDSQPDLVKPGASVKISCKASGYSTGYMHVWVKOSCKGLEWIGRINPNNGVTLY 60
Db 20 QVQLQQWAGALLKSETLSLTCGVYGGSGYYSWIRQPPCKGLEWIGIN-HSGITNY 78
QY 61 NOKFKDKATLVDSKSTTAYMELRSLTSEDSAVYYCARSTMIT-----NYVMYWGQ 112
Db 79 NPSLKSRTVISVDTSKKSLKSLSVNAADTAVYYCAR--VITRASPGTDRGMDWVGQ 136
QY 113 GTSVTVSAAKTPPSVPL--APGSAATNSMTGLGVLKGYFPFVTVW---NSGSL 167
Db 137 GTTVTVSSGSASAPTLFPLVSCENSPSDTSS-VAVGCLAQDFLPDSITFSMKYKNNSDIS 195
QY 168 SGVHTFPAVLQSDLYTLSSSVTVPPSS---TWPSSETVTCNVAHPASSTKVKKIIVPRDSCG 224
Db 196 S-TRGFFSVLRGGKYAATSOALLPKSKVMQGTDEHVCKVQHP-NGNKEKNVPLPVAIAEL 253
QY 225 PSEKSEINEKDLRKSELOGTALGNLQAIYYNSKAITSSSEKSADQFLTNTLFLKGFET 284
Db 254 PPKVSVFVPPRD-----QATGFSP-----GFF- 268
QY 285 GHPIYNDLLVGLGTAATSEYEGSSVDLYGAYGYQCAGTTPNKTACMYGGVTLHDNNRL 344
Db 269 GNPRKSLIC-----QATGFSP----- 285
QY 345 TEKKKVPINLNLWDGKQ--TTVPIDKVKTSKKE-----VTVOELDLQARHYLHGK 391
Db 286 ---ROIQVSWLRGKGVGGVTTDQVQAEAKESGPTTYKVTSTLTIKESD----- 332
QY 392 FGLYNSDSFGKGVQGLVIFHSSEGSTVSYDLFDAQGYDPTLLRIYRDNTTISTSLSI 451
Db 333 --WLSQSMFTCRVDHRLGTLFQONASSMCVDP-----QDTAIRVF-----AIPPSFASI 378
QY 452 SLYLITTSIVMTQPTSLLSVAGDRVTITCKASQSVNDVAVYQKQPSKLLISYTS 511
Db 379 FL--TKSTKLTLCLVTDL--TTYDSVTI-----SWTRQN-GEAVK---THNINI 417
QY 512 RYAGVDPDRFSSGVTGDTTLTISSVQAEADAAYFCQDYNSPPTF-----GGGKTL 562
Db 418 SESHPNATFSAVG-----EASI--CEDDWSNGERFTCTVTHTDLPSPKQ 460
QY 563 EIKRADA--PTVSIFPPSSSEOLT-SGGASVVCFLNNFYPKDINVKKWIDGS--ERQN 615
Db 461 TISRPKGVALLHRPDVYLLPAREQLNLRRESATITCLVTGSPADVFQVMQRGQPLSPK 520
QY 616 GVLSNWDQDQSKDSTYSMSSTLTLTKDYEYRHNSYTCEATHK 657
Db 521 YVTSAPMPEQAPGRYFAHSILTVSEEWNTGETYTCVVAHE 562

RESULT 34
Q8WUX4 PRELIMINARY; PRT; 588 AA.
AC Q8WUX4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 64.4 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
```



```
Db 461 TISRPGVALHRPDVYLLPPAREQLNLRRESATITCLVTGSPADVFVQWMQRPGLSPEK 520
QY 616 GVLNSMTDQDSDKSTYSMSSTLTTLTKDEYERHNSYTCETHK 657
Db 521 YVTSAPMPEQAPGRYFAHSILTVSEEWNTGETYTCVVAHE 562
RESULT 36
Q96AA6 PRELIMINARY; PRT; 618 AA.
AC Q96AA6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 67.8 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017356; AAH17356.1;
DR InterPro; IPR003598; Ig_c2.
DR pfam; PF00047; Ig; 5.
DR SMART; SM00408; Igc2; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein; Immunoglobulin domain.
SQ SEQUENCE 618 AA; 67758 MW; 96DBD4C7C696E0A6 CRC64;

Query Match 15.3%; Score 540; DB 4; Length 618;
Best Local Similarity 25.9%; Pred. No. 1.5e-25;
Matches 182; Conservative 102; Mismatches 214; Indels 204; Gaps 29;

QY 1 EVLOQSGDPLVKPGASVKISKASGYSTFTGYMHVWKQSPGKGLWIGRINPNNGVTLY 60
Db 20 QVQLQWAGALLRPSETLSITCGYSGSYWISWIRPPGKGLWIGFIN-HSGSTNY 78
QY 61 NQKFKDKATLVDSKSTTAYMELRLSTSDSAVYYCARSTMIT-----NYVMDYWGQ 112
Db 79 NPSLKRVTISVDSKQLSLKSLSSNAADTAVYYCAR--VITRASPGTDGRYGMVWGQ 136
QY 113 GTSVTVSSAKTTPSYVPL--APGSAQTNSMYTLGCLVKGYPEPTVTW---NSGSL 167
Db 137 GTTVTVSSGSASAPTLFPLVSCNSPDSSTSS--VAVGCLAQDFLPSITFSWKYKNSD 195
QY 168 SGVHTFPVAVLQSDLYTLSSSVTPSS---TWPSETVTCNVHPASSTKVDKIVPRDSGG 224
Db 196 S-TRGFPVLRGCKYAATSOVLLPSKDVMOGTDEHVCKVQHP-NGNKEKNVPLVIAEL 253
QY 225 PSKSESEINEKDLKKSELOGTALGNLKIYYNSKAITSSSEKSAQDFLNTLLFKGFFT 284
Db 254 PPKVSFVPPRD-----GFF- 268
QY 285 GHPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYCAGGTPNKTACMYGGVTLHDNNRL 344
Db 269 GNPKRKSLIC-----QATGFS- 285
QY 345 TEKKVPIINLWIDGKQ--TTVPIDKVKTSKKE-----VTQELDQARHYLHGK 391
Db 286 ---RQIQVSWLRGKQVSGVTTDQVAAKESGPTTYKVTSLTKESD----- 332
QY 392 FGLYNSDSFGKVGQRLIVFHSSEGTIVSYDLEDAQGVQPDTLRLIYRDNNTTSSLSI 451
Db 333 --WLSQSMETCRVDHRGLTFQGNASSMCPVD-----QDTAIRVF---AIPPSFASI 378
QY 452 SLYLYTTSIVMTQPTPSLLVSAGDRVTITCKASQSVNDVANYQKPGOSPKLLISYTS 511
Db 379 FL---FKSTKLTLCLVTDL--TTVDVSVTI-----SWTRQN-GEAVK---THTNI 417
```

```
QY 512 YIAGVDPDRFSGYGDFTLTLSVQAEADAAYFCQDYNPSPTF-----GGGTKL 562
Db 418 SESHPNATFSAVG-----EASI--CEDDWSGERFCTCTVTHDLPSPKQ 460
QY 563 EIKRADAA-----PTVSIFPPSSSQLT--SGGASVCFLNPFYPKDINVKWKIDGS--ERQN 615
Db 461 TISRPGVALHRPDVYLLPPAREQLNLRRESATITCLVTGSPADVFVQWMQRPGLSPEK 520
QY 616 GVLNSMTDQDSDKSTYSMSSTLTTLTKDEYERHNSYTCETHK 657
Db 521 YVTSAPMPEQAPGRYFAHSILTVSEEWNTGETYTCVVAHE 562
RESULT 37
Q8WUK1 PRELIMINARY; PRT; 613 AA.
AC Q8WUK1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 67.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020240; AAH20240.1;
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; IGc1; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671E315 CRC64;

Query Match 15.3%; Score 537.5; DB 4; Length 613;
Best Local Similarity 25.4%; Pred. No. 2.1e-25;
Matches 176; Conservative 105; Mismatches 220; Indels 193; Gaps 26;

QY 1 EVLOQSGDPLVKPGASVKISKASGYSTFTGYMHVWKQSPGKGLWIGRINPNNGVTLY 60
Db 20 QVQLVESGGGVQPGKSLRLSCAASGFTSSYGMHWVRQAPGKGLWIAVISYDGSNKYY 79
QY 61 NQKFKDKATLVDSKSTTAYMELRLSTSDSAVYYCARSTMITNYVMDYWGQTSVTSS 120
Db 80 ADSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCAKDMSEGVETFDIMGQGTMTVTS 139
QY 121 AKTTPSPVPL--APGSAQTNSMYTLGCLVKGYPEPTVTW---NSGSLSSGVHTFA 175
Db 140 GSASATPLPLVSCNSPDSSTSS--VAVGCLAQDFLPSITFSWKYKNSDISS--TRGFP 197
QY 176 VLQSDLYTLSSSVTPSS---TWPSETVTCNVHPASSTKVDKIVPRDSGGPSEKSEI 232
Db 198 VLRGGKYAATSOVLLPSKDVMOGTDEHVCKVQHP-NGNKEKNVPLVIAELPPKVSFV 256
QY 233 NEKDLRKKSELOGTALGNLKIYYNSKAITSSSEKSAQDFLNTLLFKGFFTGHFWYNDL 292
Db 257 PPRD-----GFF-GNPKRSL 271
QY 293 LVDLGSTAATSEYEGSSVDLYGAYGYCAGGTPNKTACMYGGVTLHDNNRLTEKKVPI 352
Db 272 IC-----QATGFS- 285
QY 353 NLWIDGKQ--TTVPIDKVKTSKKE-----VTQELDQARHYLHGKFLYNSDS 399
```

Db	286	SWLRCKQVSGVTTDOVAEAKESGPTTKVTSTLTIKESD-----WLSQSM	333
QY	400	FGKVGRLIVFHSRSGSVSYDLFDAQGYDPOTLLRIYRDNTTISSTLSISLYLYTTS	459
Db	334	FTCRVDHRLGTLTFOONASSMCVPD-----QDTAIRVF-----AIPSPASIFL---TKS	378
QY	460	IVNTQPTSLLSVAGDRVTITCKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVPDR	519
Db	379	TKLTLCLVTLDTL--TTYDSVTI-----SWTRQN-GEAVK---THTNISESHPNAT	420
QY	520	FSGSGVGTDTLTISSVQAEADAAYFCQDYNPPPTF-----GGCTKLEIKRADAA	570
Db	421	FSAVG-----EASI--CEDDWSNGERTCTVTHTDLPSPKQITSRPKGV	463
QY	571	-----PTVSIFPPSPSEQLT-SGGASVYVCFLLNFYPKDINVKWKIDGS--ERQNGVLNSWTD	623
Db	464	ALHRPDVYLLPPAREQLNRESATITCLVTGFSPADVFVQWMQRGQPLSPKVTTSAPWP	523
QY	624	QDSKOSTYSMSSTLTLTQDEYERHNSYTCETHK	657
Db	524	EPOAGRYFAHSILTYSEBEWNTGETYTCVVAHE	557
RESULT 38			
Q924Q9			
ID	Q924Q9	PRELIMINARY;	PRT; 145 AA.
AC	Q924Q9;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	VH186.2-D-J-C mu protein (fragment).		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6;		
RA	Kozono Y., Kozono H., Azuma T.;		
RT	"Direct Estimation of Relative Affinity by Flow Cytometry Reveals		
RT	Affinity Maturation of B Cell Antigen Receptors in Response to (4-		
RT	Hydroxy-3-Nitrophenyl)Acetyl (NP).";		
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB067791; BAB63276.1;		
DR	InterPro: IPR003006; Ig_MHC.		
DR	Pfam: PF00047; Ig; 1.		
FT	NON_TER 1		
FT	NON_TER 145		
SQ	SEQUENCE 145 AA; 16001 MW; 0F409EB09FA333D2 CRC64;		
Query Match			
Best Local Similarity 15.2%; Score 535; DB 11; Length 145;			
Matches 99; Conservative 15; Mismatches 17; Indels 0; Gaps 0;			
QY	1	EVOLQSGDPLVKPGASVKISCKASGYSTGYMHVWKSPGKLEWIGRINPNNGVTLY	60
Db	1	QVLOQPGAEVLKPGASVKLSCKASYTETSYMHVWKQRPGRGLEWIGRIDPNSGCTKY	60
QY	61	NOKFKDKATLTVDKSTTAYMELRSLTSEDSAVYYCARSTMTITNYMDYWGQGSTVTSS	120
Db	61	NEKFSKATLTVDKPSMTAYMQLSSLTSEDSAVYYCARSTMTITNYMDYWGQGSTVTSS	120
QY	121	AKTTPPSVYPL	131
Db	121	ESQSFNVFPL	131
RESULT 39			
Q91WT3			
ID	Q91WT3	PRELIMINARY;	PRT; 481 AA.
AC	Q91WT3;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		

DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Hypothetical 52.0 kDa protein.			
GN	AI893585.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=COLON;			
RA	Strausberg R.;			
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC013488; AAH13488.1; -			
DR	MGI; MGI:2144917; AI893585.			
DR	InterPro; IPR003598; Ig_C2.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	Pfam; PF00047; Ig; 4.			
DR	SMART; SM00408; IGC2; 1.			
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_2.			
KW	Hypothetical protein; Immunoglobulin domain.			
SQ	SEQUENCE 481 AA; 52022 MW; 4EEB5C253038B718 CRC64;			
Query Match				
Best Local Similarity 15.2%; Score 535; DB 11; Length 481;				
Matches 172; Conservative 79; Mismatches 168; Indels 246; Gaps 22;				
QY	1	EVOLQSGDPLVKPGASVKISCKASGYSTGYMHVWKSPGKLEWIGRINPNNGVTLY	60	
Db	20	QVLOQPGAEIVRPGASVKLSCKTSGYTFDYMHVWKQRPGRGLEWIGALDPFDSYTSY	79	
QY	61	NOKFKDKATLTVDKSTTAYMELRSLTSEDSAVYYCARSTMTITNYMDYWGQGSTVTSS	120	
Db	80	NOKFKG--TTLTVDTSSTSSAYMLLSLTSEDSAVYFCARGPRDSSGY--YWGQGTTLTVSS	136	
QY	121	AKTTPPSVYPLAPGSAATNSMTLCLVKGYEPE-PVTVTW-NSGSLSSGVHTFFAVLQ	178	
Db	137	EPAREPTIYPLTPPQALSSDPVI-IGCLIHDPSTGMNVTMGKSGKDIITVNPFPALAS	195	
QY	179	SDLYTLSSSVTPSPSTWP-SETVTCNVAPASSTKVKIVPRDSGGPSEKSEINEKDL	237	
Db	196	GGRYTHSSQLTLPAVECPGESVKCSVQH-----DSNPVQELNV---234		
QY	238	RKKSLEQGTALGNLKOIYYYNKSKAITSSSEKSAOFLNTLLFKGFTGHFWYNDLLVDLG	297	
Db	235	-----NCPGIC--SPPTTPPPSPCQPSLS-----LQRPALDILL--LG	268	
QY	298	STAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMGVGTLDHNNRLTEKKVPINLWID	357	
Db	269	SDASIT-----CTLNG-----279		
QY	358	GKQTTVPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGKVGQGLIVFHSSEGS	417	
Db	280	-----279		
QY	418	TVSYDLFDAQGYDPDTLLRIYRDNTTISSTLSISLYLYTTSIVMTQPTTSLIVSAGDRV	477	
Db	280	-----LRDPEG-----AVFTWEP-----TGKDAV	299	
QY	478	TITCKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVPDRFSGSGYGTDFTLTSSVQ	537	
Db	300	-----OKKAVQNSCGGYS-----VSSVLPGCAERWNS--GASFCTCTVTHPE	338	
QY	538	AEDAAVYFCQDYNPPPTFGGTTKLEIKRADAAPTYSIRPPSPSEQLTSGG-ASVVCFLNN	596	
Db	339	SDTLTGTTIAKVTVNTFP-----PQVHLLPPPEELALNELVSLTCLVRA	382	
QY	597	FYPKDNVWKIDGSRQNG-----VLNSWTDQDSKSTYSMSSTLTLTQDEYERHNSYTC	652	
Db	383	FNPKEVLVRW-LHGNEELSPESLYLFEPLKEPGEGATTLYTSLVRSVAEIKWQGOYSC	441	
QY	653	EATHK	657	
Db	442	MVGHE	446	

Db	20	QVHLVQSGAEVMSPGASRVRSCKTSGYAFHTYSLIYWRQAPGOGLEWGMWISPS	SNTFR	79
Qy	61	NOKEFKRATLTVOKSSTAYAMELRSLTSEDSAVVYCAR	S---	112
Db	80	AKKFOGRVTLTDTSTVYAMELRSRSDTAYVYCARRYC	SYSSCONDY	139
Qy	113	GTSVTVSSAKTTPSPVYPLAPGSAAOQNSMWTGLCVK	GYFP-EPVTV	171
Db	140	GTTVTWSSASPTSPKVPPLSLCS--TOPDGNWVITACL	VOGFFPQEP	198
Qy	172	TFPAV--LQSDLYTLSSSVIVPSS-TWPSETVTCNAH	---PASSTKV	225
Db	199	NFPSPQASGDLTYTTSQLTLPATQCLAGKSVTCHVKH	YTNPSQDVT	257
Qy	226	SEKSEETNEKDLARKSELOQTALGNLKOIYYNSKAIT	SEKSAQOFL	285
Db	258	SP-----STPTTSPSCCHPRLS-----	-----LH	277
Qy	286	HPWYNLDVLGLGTAATSEYEGSSVDLYGAYGYQCAG	GTPTNACMYG	345
Db	278	RPALDEL--LGSEA-----	-----NLCTLTG	298
Qy	346	EKKVPIINLWDGKQTVPIDKVKTSKEVTVQELDLQAR	HYLHGFEGL	405
Db	299	-----	-----	298
Qy	406	RGLIVFHSSEGSTVSYDLFDAQQGYQDTLLRIYRDN	TNTTSLSLSIS	465
Db	299	-----LRASG-----	-----VFVTT	310
Qy	466	PT---SLLVSAGDRVTTITKASOSVSNDA--WYQKQ	PGQSPKLLIS	520
Db	311	PSSGKSAVQPPDRDLCCGYSVSVLSGCAEPW--	---NHGKTFTCT	359
Qy	521	SGSGYGTDFLTLTSSVQDAEAAYFCQODYNSPPT	GGGTGLEIKR	580
Db	360	-----TPLTATUS-----	-----KSGNTE-----	384
Qy	581	EQLTSGG--ASVVCFLNNFYPKDINVKWKIDGSR	QRQGNVLSW	636
Db	385	EELALNELVILTCLARGFSFKDVLVRHMQSQEL	PREKYLTVAS	444
Qy	637	LTLTKDYEYRHNSVYTCATHK	657	
Db	445	LRVAEDWKKGDTFSCVMGHE	465	

```

Best Local Similarity 74.2%; Pred. NO. 4.7e-25;
Matches 98; Conservative 15; Mismatches 18; Indels 1; Gaps 1;

QY 1 EVQLQSGPDLVKPCASVKISCKASGYSTGYMHMWVKQSPGKGLIEWIGRINPNNGVTLY 60
      :||||| : |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 QVQLQQPAELVKPCASVKLSCKASGYTFTSYMHMWVKQRPGRGLIEWIGRIDPNSSGGTKY 60
      :||||| : |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NQKFKDKATLTVDKSTTAYMELRLSLTSDSAVYYCARSTM-ITNYVMYDYGQGSVTVS 119
      :||||| : |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NEKFSKATLTVDKFSSTAYMQLSLTSDSAVYYCARSLYDYGDAYMDYMGQGSVTVS 120
      :||||| : |||||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 120 SAKTTPPSVYPL 131
      | : |:|||
Db 121 SESQSFPNVFPL 132
      : |||||

RESULT 47
Q924P7 PRELIMINARY; PRT; 145 AA.
AC Q924P7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE VH186.2-D-J-C mu protein (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB065918; BAB63934.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 15988 MW; FB73958704796C9A CRC64;

Query Match 14.6%; Score 515; DB 11; Length 143;
Best Local Similarity 72.5%; Pred. No. 7.7e-25;
Matches 95; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYMHVWKSPGKGLWIGRINPNGVTLY 60
Db 1 QVLOQPGALVPGSSVKLSCKASGYFTSYMHVWKORPGKGLWIGRIDPNSGTTY 60
QY 61 NQKFKDKATLTVDKSSTTAYMELRLSITSDSAVYYCARSTMTITNYVMDYWGQTSVTSS 120
Db 61 NQKFKSKATLTVDKPSSTAYMQLSITSDSAVYYCARSGSFYAMDYWGQTSVTSS 120
QY 121 AKTTPPSVYPL 131
Db 121 ESQSFNPVPL 131

RESULT 48
Q924R0 PRELIMINARY; PRT; 143 AA.
AC Q924R0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)
DE VH186.2-D-J-C mu protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB067790; BAB63275.1; -.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00408; Igc2; 1.
KW Immunoglobulin domain.
FT NON_TER 1
FT NON_TER 143
SQ SEQUENCE 143 AA; 15868 MW; 139B2E9668B1E07F CRC64;

Query Match 14.6%; Score 514; DB 11; Length 143;
Best Local Similarity 74.0%; Pred. No. 8.7e-25;
Matches 97; Conservative 15; Mismatches 17; Indels 2; Gaps 1;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYMHVWKSPGKGLWIGRINPNGVTLY 60
Db 1 QVLOQPGALVPGASVKLSCKASGYFTSYMHVWKORPGKGLWIGRIDPNSGTTY 60
QY 61 NQKFKDKATLTVDKSSTTAYMELRLSITSDSAVYYCARSTMTITNYVMDYWGQTSVTSS 120
```

```
Db 61 NQKFKSKATLTVDKPSSTAYMQLSITSDSAVYYCAR--WDEDYAMDYWGQTSVTSS 118
QY 121 AKTTPPSVYPL 131
Db 119 ESQSFNPVPL 129

RESULT 49
Q8WU38 PRELIMINARY; PRT; 573 AA.
AC Q8WU38;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 63.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021276; AAH21276.1; -.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 4.
DR SMART: SM00409; Ig; 1.
DR SMART: SM00407; Igc1; 3.
DR SMART: SM00406; Igv; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 573 AA; 62967 MW; FD072344033AC530 CRC64;

Query Match 14.6%; Score 514; DB 4; Length 573;
Best Local Similarity 27.8%; Pred. No. 5.6e-24;
Matches 157; Conservative 85; Mismatches 201; Indels 122; Gaps 18;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYMHVWKSPGKGLWIGRINPNGVTLY 60
Db 20 EVLVESGGGLVQGRSLRLSCAASGFTFDYAMHWVRQAPGKGLWVSGISNNGSIGY 79
QY 61 NQKFKDKATLTVDKSSTTAYMELRLSITSDSAVYYCAR----STMTITNYVMDYWGQTSV 116
Db 80 ADSVKGRFTISRDNKNSLYLQMSLRAEDTALYYCAKHGSGSYIGYYGMVWGQTTV 139
QY 117 TVSSAKTTPSVYPLAPGSA-AQTNSMVTLGCLVKGYFPEPTVTWNSGSLSCVHTFPA 175
Db 140 TVSSAPTKAPDVEPTIISGCRHPKDNSPVVLACLTIGYHPTSVTVTWMTGOSQORTFPE 199
QY 176 VLQSDTLTSSS-VTVPSSTPSETVTCNVAPAS-----STKVDKKIVP---- 219
Db 200 IQRRDSYMTSSQLSTPLQWRQGEYKCVVQHTASKSKKEIFRPESPKAQASVPPTAQP 259
QY 220 -----RDSGGPSEKSEEEINEKDLRKKSLEQGTALGNLKI--YYNSKA 261
Db 260 QAEGSLAKATTAPATTATNTGRGGEKKKEKEQEERETKTPECPSHTQPLGVYLLTPA 319
QY 262 ITSSEKSAQOFLNTLLFKGFFTHPWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYQC 321
Db 320 V-----QDLWLRLDKATFTCFVYVGS-----DLKDAHLTWEV 349
QY 322 AGGTPTNKACMGV-----TLHDNNRLTEKK--VPINLWIDGKQTVPIDVKV-TSKK 373
Db 350 AGKVPT-----GGVEGLLERHSNGSQSHSLTLPRLSNAGTSVTCTLNHPSLPQR 403
QY 374 EVTVQELDLQARHYLHGKFGLYNSDFGGKVGQRLGVFIHSS--GSTVSYDLFDAQGGYPD 432
Db 404 LMALEPAAQA-----PVKLSLNLASSDPPPEAASWLCVSGSFSP 445
```

QY 433 TLRIY-RDNTTSSLSI-----SLYLTYSIVMTQTPTSLVSAGDRVITICKA 483
Db 446 NILLMLEQREVTSGFAPRPPQPGSTTFWAWSVLRVPAPPS-----PQPATYTCVV 500
QY 484 SQSVNDVAVYQKPGQSPKLLISY 508
Db 501 SHEDSRLL-----NARSLEVS 519

RESULT 50

Q9QXE9 PRELIMINARY; PRT; 117 AA.
AC Q9QXE9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Immunoglobulin heavy chain V-D-J region (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP Clemen A., Rademaekers A., Specht C., Koelsch E.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ225174; CAB65237.1; -
DR HSSP; P01810; 2FBJ
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13000 MW; CDDE2AF84D499734 CRC64;

Query Match 14.4%; Score 508.5; DB 11; Length 117;
Best Local Similarity 82.5%; Pred. No. 1.5e-24;
Matches 99; Conservative 5; Mismatches 13; Indels 3; Gaps 1;
QY 1 EVOLQSGPDLVKPGASVKISKASGYSTGYMHWVKSPGKLEWIGRINPNNGTLY 60
Db 1 EVOLQSGPELVKPGASVKMSCKASGYTDTYMKWVKSHGKSLWIGDINPNNGTYS 60
QY 61 NOKFKDKATLTVDKSSSTAYMELSLTSEDSAVYVCARSTMTNTNYMDYWGQTSVTSS 120
Db 61 NOKFKDKATLTVDKSSSTAYMELSLTSEDSAVYVCARDRY---YAMDYWGQTSVTSS 117

Search completed: January 8, 2003, 11:58:21
Job time : 55.5525 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:52:01 ; Search time 16.9923 Seconds
(without alignments)
1827.149 Million cell updates/sec

Title: US-09-900-766-2

Perfect score: 1218

Sequence: 1 SEKSEINEKDKRKSELOG.....RDNTTISSTLSISLYLTT 233

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID52/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID52/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID52/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID52/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID52/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID52/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID52/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID52/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID52/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID52/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID52/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID52/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID52/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID52/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID52/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID52/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID52/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID52/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID52/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID52/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID52/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID52/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID52/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1113	91.4	245	18	Staphylococcus ent
2	1107	90.9	257	22	Peptide sequence f
3	1093	89.7	230	22	Staphylococcus aur
4	1091	89.6	245	18	Staphylococcus ent
5	1048	86.0	230	14	Staphylococcus ent
6	1044	85.7	230	23	Staphylococcus aur
7	1035	85.0	230	12	Staphylococcus ent
8	960	78.8	233	12	Staphylococcus ent
9	948	77.8	257	22	Peptide sequence f
10	944	77.5	233	18	Staphylococcus ent

11	942	77.3	233	14	AA45011	Staphylococcal ent
12	941	77.3	233	18	AAW35373	Staphylococcus ent
13	941	77.3	233	22	AAW67338	Staphylococcus aur
14	925	76.6	233	23	ABB76234	Staphylococcus aur
15	925	75.9	257	23	ABB79501	Staphylococcal ent
16	921	75.6	233	21	AA544463	Amino acid sequenc
17	921	75.6	233	23	ABB79502	Staphylococcal ent
18	908	74.5	257	21	AAW70102	Staphylococcal ent
19	903	74.1	233	21	AAW70103	Mutant Staphylococ
20	605	49.7	228	14	AA45013	Staphylococcal ent
21	605	49.7	228	22	AAW67340	Staphylococcus aur
22	605	49.7	228	23	ABB76236	Staphylococcus aur
23	597	49.0	228	12	AAW13205	Staphylococcal ent
24	370	30.4	91	18	AAW24299	Staphylococcus aur
25	291.5	23.9	259	23	ABP29357	Streptococcus poly
26	287.5	23.6	239	20	AAV06254	Staphylococcal gro
27	287.5	23.6	239	20	AAV06253	Staphylococcal gro
28	282.5	23.2	239	20	AAV06256	Staphylococcal gro
29	280.5	23.0	239	20	AAV06255	Staphylococcal gro
30	279.5	22.9	239	20	AAV06252	Staphylococcal gro
31	277.5	22.8	238	14	AA45016	Staphylococcal ent
32	275.5	22.6	238	22	AAW67343	Staphylococcus aur
33	275.5	22.6	238	23	ABB76239	Staphylococcus aur
34	273.5	22.5	228	22	AAW63856	Amino acid sequenc
35	272.5	22.4	238	12	AAW13208	Staphylococcal ent
36	271.5	22.3	251	18	AAW12153	Streptococcus pyog
37	270.5	22.2	221	12	AAW13209	Streptococcus pyog
38	270.5	22.2	221	14	AAW45017	Staphylococcal ent
39	270.5	22.2	221	23	ABB76240	Staphylococcus pyo
40	270.5	22.2	251	18	AAW12150	Streptococcus pyog
41	270.5	22.2	251	18	AAW12151	Streptococcus pyog
42	268.5	22.0	251	18	AAW12154	Streptococcus pyog
43	268.5	22.0	251	18	AAW12148	Streptococcus pyog
44	268.5	22.0	251	18	AAW12147	Streptococcus pyog
45	268.5	22.0	251	18	AAW12146	Streptococcus pyog

ALIGNMENTS

RESULT 1
AAW35375
ID AA35375 standard; peptide: 245 AA.
XX
AC AA35375;
XX
DT 20-APR-1998 (first entry)
XX
DE Staphylococcus enterotoxin SEE modified superantigen.
XX
KW SEE; SEA; Staphylococcus enterotoxin; superantigen; conjugate;
XX treatment; cancer; infection; autoimmune disease; antibody; modified.
XX
OS Staphylococcus sp.
FH Key Location/Qualifiers
FT Misc-difference 20 /label= R20G
FT /note= "wild-type Arg is replaced by Gly"
FT Misc-difference 21 /label= N21T
FT /note= "wild-type Asn is replaced by Thr"
FT Misc-difference 24 /label= S24G
FT /note= "wild-type Ser is replaced by Gly"
FT Misc-difference 27 /label= R27K
FT /note= "wild-type Arg is replaced by Lys"
PN W09736932-A1.
XX
PD 09-OCT-1997.
XX

AAW35374 standard; peptide; 245 AA.
AAW35374;
20-APR-1998 (first entry)
Staphylococcus enterotoxin SEE wild-type superantigen.
SEE; SEA; Staphylococcus enterotoxin; superantigen; conjugate;
treatment; cancer; infection; autoimmune disease; antibody.
Staphylococcus sp.
Key Location/Qualifiers
FT Misc-difference 20 /note= "can be mutated at this position"
FT Misc-difference 21 /note= "can be mutated at this position"
FT Misc-difference 24 /note= "can be mutated at this position"
FT Misc-difference 27 /note= "can be mutated at this position"
WO9736932-A1.
09-OCT-1997.
26-MAR-1997; 97WO-SE00537.
12-AUG-1996; 96US-0695692.
29-MAR-1996; 96SE-0001245.
(PHAA) PHARMACIA & UPJOHN AB.
Abrahmsen L, Antonsson P, Bjoerk P, Dohlsten M;
Forsberg G, Hansson J, Kalland T;
WPI; 1997-503052/46.
Conjugate of target seeking moiety and modified superantigen -
useful for activating the immune system to treat cancer, viral
infections, parasitic infestations and autoimmune diseases
Claim 4; Pages 38-39; 58pp; English.
This is the wild-type Staphylococcus enterotoxin SEE superantigen. This
SEE superantigen can be modified to be used in a novel conjugate. The
novel conjugate comprises a target seeking moiety and a modified wild
type superantigen. The modified superantigen retains its ability to
activate a subset of T cells, even though 1 or more wild-type amino acid
residues in at least 1 region which functions in determining binding to
T cell receptor (TCR) and activation of a subset of T cells has/have been
replaced. Such a modified superantigen can optionally be used as part of
a conjugate with a target seeking moiety, for activating the immune
system to treat a mammalian disease. A pharmaceutical composition can be
prepared comprising a modified antibody (preferably a Fab fragment fused
to a peptide moiety providing activation of T cells in vbeta specific
manner) in which cysteines providing for interchain cysteine linkages in
the native antibody have been replaced (preferably by serine residues) to
prohibit cysteine formation. The modified wild-type superantigen is used
for treating cancer, viral infections, parasitic infestations and
autoimmune disease. The modified wild type superantigen has a lower
immunogenicity and reactivity with neutralising antibodies and has fewer
side-effects when used as a drug, compared to wild type superantigen.

Query Match 89.7%; Score 1093; DB 22; Length 230;
Best Local Similarity 89.6%; Pred. No. 6,7e-103;
Matches 206; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 4 SEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSAOQFTNTLLFKGFTGHPW 63
DB 1 SEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSAOQFTNTLLFKGFTGHPW 60
QY 64 YNDLLVDLGSSTAEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLTEK 123
DB 61 YNDLLVDLGSSTAEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLTEK 120
QY 124 KVPINLWIDGKQTPVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRL 183
DB 121 KVPINLWIDGKQTPVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRL 180
QY 184 IVFHSSEGSTVSYDLFDAGQGVPTDILLRIYRDNWTTSTSLISLYLTT 233
DB 181 IVFHSSEGSTVSYDLFDAGQGVPTDILLRIYRDNWTTSTSLISLYLTT 230

RESULT 4
AAW35374

ID AAW35374 standard; peptide; 245 AA.
XX AC AAW35374;
XX DT 20-APR-1998 (first entry)
XX DE Staphylococcus enterotoxin SEE wild-type superantigen.
XX KW SEE; SEA; Staphylococcus enterotoxin; superantigen; conjugate;
XX OS treatment; cancer; infection; autoimmune disease; antibody.
XX OS Staphylococcus sp.
XX FH Key Location/Qualifiers
FT Misc-difference 20 /note= "can be mutated at this position"
FT Misc-difference 21 /note= "can be mutated at this position"
FT Misc-difference 24 /note= "can be mutated at this position"
FT Misc-difference 27 /note= "can be mutated at this position"
XX PN WO9736932-A1.
XX PD 09-OCT-1997.
XX PF 26-MAR-1997; 97WO-SE00537.
XX PR 12-AUG-1996; 96US-0695692.
XX PR 29-MAR-1996; 96SE-0001245.
XX PA (PHAA) PHARMACIA & UPJOHN AB.
XX PI Abrahmsen L, Antonsson P, Bjoerk P, Dohlsten M;
XX PI Forsberg G, Hansson J, Kalland T;
XX DR WPI; 1997-503052/46.
XX PT Conjugate of target seeking moiety and modified superantigen -
XX PT useful for activating the immune system to treat cancer, viral
XX PT infections, parasitic infestations and autoimmune diseases
XX PS Claim 4; Pages 38-39; 58pp; English.
XX CC This is the wild-type Staphylococcus enterotoxin SEE superantigen. This
XX CC SEE superantigen can be modified to be used in a novel conjugate. The
XX CC novel conjugate comprises a target seeking moiety and a modified wild
XX CC type superantigen. The modified superantigen retains its ability to
XX CC activate a subset of T cells, even though 1 or more wild-type amino acid
XX CC residues in at least 1 region which functions in determining binding to
XX CC T cell receptor (TCR) and activation of a subset of T cells has/have been
XX CC replaced. Such a modified superantigen can optionally be used as part of
XX CC a conjugate with a target seeking moiety, for activating the immune
XX CC system to treat a mammalian disease. A pharmaceutical composition can be
XX CC prepared comprising a modified antibody (preferably a Fab fragment fused
XX CC to a peptide moiety providing activation of T cells in vbeta specific
XX CC manner) in which cysteines providing for interchain cysteine linkages in
XX CC the native antibody have been replaced (preferably by serine residues) to
XX CC prohibit cysteine formation. The modified wild-type superantigen is used
XX CC for treating cancer, viral infections, parasitic infestations and
XX CC autoimmune disease. The modified wild type superantigen has a lower
XX CC immunogenicity and reactivity with neutralising antibodies and has fewer
XX CC side-effects when used as a drug, compared to wild type superantigen.
XX SQ Sequence 245 AA;
Query Match 89.6%; Score 1091; DB 18;
Best Local Similarity 85.3%; Pred. No. 1.2e-102;
Matches 209; Conservative 9; Mismatches 15; Indels 12; Gaps 1;
QY 1 SEKSEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSAOQFTNTLLFKGFTG 60
|||||
|||||

Db 1 SEKSEINEKDLRKKSELQARNALSLRQIYYNEKAITENKESDDQFLNTLLFKGFFTG 60
Qy 61 HPWYNLLVDLGSTAAATSEYEGSSVDLYCAYYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 61 HPWYNLLVDLGSKDATNKYKGGKVDLYGAYYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
Qy 121 EEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHVLHGKFGLYNSDSFGGKVQ 180
Db 121 EEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHVLHGKFGLYNSDSFGGKVQ 180
Qy 181 RGLIVFHSSEGSTVSVDLFDAGQGYPTDRLRIYRDNRT-----TISSTSLSL 228
Db 181 RGLIVFHSSEGSTVSVDLFDAGQGYPTDRLRIYRDNKTINSENHLIDLYTINSENHLIDL 240
Qy 229 YLYTT 233
Db 241 YLYTT 245
RESULT 5
ID AAR45012 standard; protein; 230 AA.
XX AAR45012;
XX AAR45012;
DT 08-JUN-1994 (first entry)
XX Staphylococcal enterotoxin SEE.
DE Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
KW auto-immune disease; toxicity; Protein A; perfusion system.
XX Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 120 /note= "Given in the specification as J, no further
FT details given"
FT Misc-difference 121 /note= "Given in the specification as J, no further
FT details given"
FT Misc-difference 123 /note= "Given in the specification as O, no further
FT details given"
FT Misc-difference 124 /note= "Given in the specification as U, no further
FT details given"
PN W09324136-A.
XX
XX
PD 09-DEC-1993.
XX
PF 01-JUN-1993; 93WO-US05213.
XX
PR 01-JUN-1992; 92US-0891718.
XX
PA (STON/) STONE J L.
XX
PA (TERM/) TERMAN D S.
XX
PI Stone JL, Terman DS;
XX
XX WPI; 1993-405418/50.
XX
PT Use of staphylococcal enterotoxin(s) and homologues - for
PT treating cancer in a patient or for the treatment of auto-immune
PT diseases
XX
PS Disclosure; Fig 1; 90pp; English.
XX
CC The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
CC which may be used in the methods of the invention for treating cancer
CC in a patient. These SEs, and homologues of them, can be used as
CC tumouricidal agents for treating cancers and autoimmune disease.

CC They exhibit tumouricidal activity and toxicity identical to that
CC observed for the Protein A perfusion system. They may be administered
CC by i.v. injection.
XX
SQ Sequence 230 AA;
Query Match 86.0%; Score 1048; DB 14; Length 230;
Best Local Similarity 85.7%; Pred. No. 2.6e-98;
Matches 197; Conservative 12; Mismatches 21; Indels 0; Gaps 0;
Qy 4 SEEINEKDLRKKSELQALGNLKOIYYNSKAITSEKSAQDQFLNTLLFKGFFTGHPW 63
Db 1 SEEINEKDLRKKSELQARNALSLRQIYYNEKAITENKESDDQFLNTLLFKGFFTGHPW 60
Qy 64 YNDLLVDLGSTAAATSEYEGSSVDLYCAYYGYOCAGGTPNKTACMYGGVTLHDNNRLTEK 123
Db 61 YNDLLVDLGSKDATNKYKGGKVDLYGAYYGYOCAGGTPNKTACMYGGVTLHDNNRLTEK 120
Qy 124 KVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHVLHGKFGLYNSDSFGGKVQ 183
Db 121 VXXBKWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHVLHGKFGLYNSDSFGGKVQ 180
Qy 184 IVFHSSEGSTVSVDLFDAGQGYPTDRLRIYRDNRTTISSTSLSLYTT 233
Db 181 IVFHSSEGSTVSVDLFDAGQGYPTDRLRIYRDNKTINSENHIDIYLYTT 230
RESULT 6
ABB76235
ID ABB76235 standard; protein; 230 AA.
XX ABB76235;
XX ABB76235;
DT 09-AUG-2002 (first entry)
XX Staphylococcus aureus enterotoxin E.
DE Enterotoxin E; SEE; superantigen; antigen; tumour; cancer;
KW antitumour; therapy.
XX Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 120 /note= "given as 'J' in the specification"
FT Misc-difference 121 /note= "given as 'J' in the specification"
FT Misc-difference 123 /note= "given as 'O' in the specification"
FT Misc-difference 124 /note= "given as 'U' in the specification"
FT Misc-difference 125 /note= "given as 'V' in the specification"
XX
XX US2002051765-A1.
XX
XX 02-MAY-2002.
XX
XX 19-DEC-2000; 2000US-0741503.
XX
XX 31-JAN-1994; 94US-0189424.
XX 19-JUN-1995; 95US-0491746.
XX 03-OCT-1989; 89US-0416530.
XX 17-JAN-1990; 90US-0466577.
XX 17-JAN-1991; 91WO-US00342.
XX 01-JUN-1992; 92US-0891718.
XX 02-MAR-1993; 93US-0025144.
XX
XX (TERM/) TERMAN D S.
XX
XX Terman DS;
XX
XX WPI; 2002-415198/44.
DR

XX Reagent for treating cancer without the need for e.g. radiotherapy,
 PT comprises a specific V beta subset of T cells sensitized to a growing
 PT tumor and stimulated with superantigens -
 XX Disclosure; Fig 2; 17pp; English.

XX The present sequence is the protein sequence of enterotoxin E
 CC (SEE) of *Staphylococcus aureus*. Similarity is shown, in several
 CC stretches of sequence, between staphylococcal enterotoxins,
 CC streptococcal pyrogenic exotoxins and staphylococcal exfoliative
 CC toxins (see ABB76234-44). In the present invention, synthetic
 CC polypeptides useful in tumour therapy and in blocking or destroying
 CC autoreactive T and B lymphocyte populations are characterised by
 CC substantial structural homology to staphylococcal enterotoxin A and
 CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with
 CC statistically significant sequence homology and similarity (Z value
 CC of Lipman and Pearson algorithm in Monte Carlo analysis exceeding
 CC 6) to include alignment of cysteine residues and similar hydrophathy
 CC profiles. These superantigens are used to treat solid tumours,
 CC including their metastases, without radiation, surgery or
 CC standard chemotherapeutic agents. A claimed method of human cancer
 CC treatment involves contacting haematopoietic cells from a patient
 CC with one or more superantigens ex vivo to generate stimulated cells,
 CC selecting a specific V beta subset of cells, and reintroducing
 CC these cells into the patient to induce an in vivo therapeutic,
 CC tumouricidal reaction.

XX Sequence 230 AA;

Query Match 85.7%; Score 1044; DB 23; Length 230;
 Best Local Similarity 85.7%; Pred. No. 6.5e-98;
 Matches 197; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

Qy 4 SPEINEKDLRKSELOGTALGNLKOIYYNSKAITSSSEKSDQFLTNTLLPKGFTGHPW 63
 Db 1 SEEINEKDLRKSELRNALNLROIYYNEKAITENKESHDDQFLQHTILFKGFTGHPW 60

Qy 64 YNDLLVDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEEK 123
 Db 61 YNDLLVDGSKDATNKYKGGKVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEEX 120

Qy 124 KVPINLWDGKQTTVPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRGL 183
 Db 121 XVXXXKWDGKQTTVPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRGL 180

Qy 184 IVFHSSEGSTVSYDLFDAQQGQYPTDLLRIYRDNTTISSTLSISLYLTT 233
 Db 181 IVFHSSEGSTVSYDLFDAQQGQYPTDLLRIYRDNTTINSENHIDIYLYTT 230

RESULT 7
 AAR13204
 ID AAR13204 standard; Protein; 230 AA.
 XX AAR13204;
 AC AAR13204;
 XX 15-OCT-1991 (first entry)
 DT Staphylococcal enterotoxin E.
 DE Staphylococcus aureus.
 XX SEE; cancer treatment; pyrogen; tumouricide.
 KW Staphylococcus aureus.
 OS WO9110680-A.
 PN 25-JUL-1991.
 PD 17-JAN-1991; 91WO-US00342.
 XX 17-JAN-1991; 91WO-US00342.
 PF 17-JAN-1990; 90US-0466577.
 PR 17-JAN-1990; 90US-0466577.
 XX

PA (TERM/) TERMAN D S.
 XX Terman DS;
 PI WPI; 1991-237984/32.
 DR
 XX Treating cancer with enterotoxin from *Staphylococcus aureus* -
 PT administered by IV injection, having same tumoricidal activity
 PT as Staphylococcal protein A without potential toxic reactions
 XX Disclosure; Fig 1; 74pp; English.

XX SEE was isolated and purified from *S. aureus*. It can be used for
 CC treating cancer, activating cytokine mediators and procoagulant
 CC systems, augmenting natural killer cell cytotoxicity, etc. The
 CC enterotoxin can be administered intravenously, optionally with
 CC ibuprofen to attenuate toxic reaction to SEE. Synthetic
 CC polypeptides having structural homology to Staphylococcal exotoxins
 CC are claimed, provided the homology includes statistically
 CC significant sequence homology, alignment of Cysteine residues and
 CC similar hydrophathy profiles.
 CC See AAR13203-R13211.

XX Sequence 230 AA;

Query Match 85.0%; Score 1035; DB 12; Length 230;
 Best Local Similarity 84.3%; Pred. No. 5.4e-97;
 Matches 194; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

Qy 4 SEEINEKDLRKSELOGTALGNLKOIYYNSKAITSSSEKSDQFLTNTLLPKGFTGHPW 63
 Db 1 SEEINEKDLRKSELRNALNLROIYYNEKAITENKESHDDQFLQHTILFKGFTGHPW 60

Qy 64 YNDLLVDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEEK 123
 Db 61 YNDLLVDGSKDATNKYKGGKVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEEX 120

Qy 124 KVPINLWDGKQTTVPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRGL 183
 Db 121 XVQXBKWDGKQTTVPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRGL 180

Qy 184 IVFHSSEGSTVSYDLFDAQQGQYPTDLLRIYRDNTTISSTLSISLYLTT 233
 Db 181 IVFHSSEGSTVSYDLFDAQQGQYPTDLLRIYRDNTTINSENHIDIYLYTT 230

RESULT 8
 AAR13203
 ID AAR13203 standard; protein; 233 AA.
 XX AAR13203;
 AC AAR13203;
 XX 15-OCT-1991 (first entry)
 DT Staphylococcal enterotoxin A.
 DE SEA; cancer treatment; pyrogen; tumouricide.
 KW Staphylococcus aureus.
 OS WO9110680-A.
 PN 25-JUL-1991.
 PD 17-JAN-1991; 91WO-US00342.
 XX 17-JAN-1991; 91WO-US00342.
 PF 17-JAN-1990; 90US-0466577.
 PR (TERM/) TERMAN D S.
 PA Terman DS;
 PI WPI; 1991-237984/32.
 DR

```
XX Treating cancer with enterotoxin from Staphylococcus aureus -
PT administered by IV injection, having some tumoricidal activity
PT as Staphylococcal protein A without potential toxic reactions
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC SEA was isolated and purified from S.aureus. It can be used for
CC treating cancer, activating cytokine mediators and procoagulant
CC systems, augmenting natural killer cell cytotoxicity, etc. The
CC enterotoxin can be administered intravenously, optionally with
CC ibuprofen to attenuate toxic reaction to SEA. Synthetic
CC polypeptides having structural homology to Staphylococcal
CC are claimed, provided the homology includes statistically
CC significant sequence homology, alignment of Cysteine residues and
CC similar hydrophathy profiles.
CC See also AAR13204-R13211.
XX
SQ Sequence 233 AA;
Query Match 78.8%; Score 960; DB 12; Length 233;
Best Local Similarity 77.3%; Pred. No. 2.4e-89;
Matches 180; Conservative 21; Mismatches 32; Indels 0; Gaps 0;
Qy 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNYSKAITSSSEKSADQFLTNTLLFKGFFTG 60
Db 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNYSKAITSSSEKSADQFLTNTLLFKGFFTG 60
Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 61 HSWYNDLLVDLGSKDIDVYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Qy 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHVLHGKFGLYNSDSFGGKVQ 180
Db 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHVLHGKFGLYNSDSFGGKVQ 180
Qy 181 RGLIVFHSSEGSTVSYDLFDAAGQYPTDILLRIYRDNNTTSSLSLSISLYTT 233
Db 181 RGLIVFHTSTEPSVNYDLFDAAGQYPTDILLRIYRDNNTTSSLSLSISLYTT 233
RESULT 9
AAU14104
ID AAU14104 standard; peptide: 257 AA.
XX
AC AAU14104;
XX
DT 21-NOV-2001 (first entry)
XX
DE Peptide sequence from Staphylococcus aureus enterotoxin A.
XX
KW Anti-retroviral; DP178-like; DP107-like; enterotoxin A;
KW antifusogenic; antiviral; HIV transmission.
XX
OS Staphylococcus aureus.
XX
PN WO200151673-A2.
XX
PD 19-JUL-2001.
XX
PF 05-JUL-2000; 2000WO-US35727.
XX
PR 09-JUL-1999; 99US-0350841.
XX
PA (TRIM-) TRIMERIS INC.
XX
PI Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
XX
XX WPI; 2001-442157/47.
XX
PT Identifying a compound that inhibits the formation of or disrupts a
PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral
PT or intracellular modulatory activity, by detecting the formation of a
```

```
PT DP107/DP178 complex -
XX
PS Disclosure; Fig 42; 259pp; English.
XX
CC The present invention relates to peptides which exhibit anti-retroviral
CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
CC DP178-like and DP107-like peptides. The DP178 peptide corresponds
CC to amino acids 639-673 of the transmembrane protein gp41 from human
CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide
CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention
CC also relates to a method of identifying compounds that inhibit the
CC formation of or disrupts a DP107/DP178 complex. The method comprises
CC detecting the formation of a DP107/DP178 complex, both in the presence
CC or absence of a test compound, in a reaction mixture containing DP107
CC and DP178 peptides. The method is useful for identifying compounds,
CC including small molecule compounds, which may themselves exhibit
CC antifusogenic, antiviral or intracellular modulatory activity. The
CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
CC retroviral, particularly HIV, transmission to uninfected cells. The
CC present sequence represents a peptide sequence from Staphylococcus aureus
CC enterotoxin A.
XX
SQ Sequence 257 AA;
Query Match 77.8%; Score 948; DB 22; Length 257;
Best Local Similarity 76.4%; Pred. No. 4.6e-88;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;
Qy 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNYSKAITSSSEKSADQFLTNTLLFKGFFTG 60
Db 25 SEKSEINEKDLRKSELOGTALGNLKOIYYNYSKAITSSSEKSADQFLTNTLLFKGFFTG 84
Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 85 HSWYNDLLVDLDFDSDIDVYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
Qy 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHVLHGKFGLYNSDSFGGKVQ 180
Db 145 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHVLHGKFGLYNSDSFGGKVQ 204
Qy 181 RGLIVFHSSEGSTVSYDLFDAAGQYPTDILLRIYRDNNTTSSLSLSISLYTT 233
Db 205 RGLIVFHTSTEPSVNYDLFDAAGQYPTDILLRIYRDNNTTSSLSLSISLYTT 257
RESULT 10
AAW06738
ID AAW06738 standard; Protein; 233 AA.
XX
AC AAW06738;
XX
DT 08-MAR-1997 (first entry)
XX
DE Staphylococcus enterotoxin A.
XX
KW Enterotoxin A; superantigen; antigen; cytokine; chemokine; T cell;
KW lymphocyte; monocyte; natural killer cell; gene therapy; cancer;
KW vaccine; adjuvant.
XX
OS Staphylococcus sp.
XX
PN WO9636366-A1.
XX
PD 21-NOV-1996.
XX
PF 20-MAY-1996; 96WO-US07432.
XX
PR 29-DEC-1995; 95US-0580806.
XX
PR 18-MAY-1995; 95US-0446918.
XX
PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY MED.
XX
PI Dow SW, Elmslie RE, Potter TA;
```


DR WPI; 1997-503052/46.
 XX Conjugate of target seeking moiety and modified superantigen -
 PT useful for activating the immune system to treat cancer, viral
 PT infections, parasitic infestations and autoimmune diseases
 XX
 XX Claim 8; Pages 36-37; 58pp; English.
 PS
 XX This is the wild-type Staphylococcus enterotoxin SEA superantigen. This
 CC SEA superantigen can be modified to be used in a novel conjugate. The
 CC novel conjugate comprises a target seeking moiety and a modified wild
 CC type superantigen. The modified superantigen retains its ability to
 CC activate a subset of T cells, even though 1 or more wild-type amino acid
 CC residues in at least 1 region which functions in determining binding to
 CC T cell receptor (TCR) and activation of a subset of T cells has/have been
 CC replaced. Such a modified superantigen can optionally be used as part of
 CC a conjugate with a target seeking moiety, for activating the immune
 CC system to treat a mammalian disease. A pharmaceutical composition can be
 CC prepared comprising a modified antibody (preferably a Fab fragment fused
 CC to a peptide moiety providing activation of T cells in Vbeta specific
 CC manner) in which cysteines providing for interchain cysteine linkages in
 CC the native antibody have been replaced (preferably by serine residues) to
 CC prohibit cysteine formation. The modified wild-type superantigen is used
 CC for treating cancer, viral infections, parasitic infestations and
 CC autoimmune disease. The modified wild type superantigen has a lower
 CC immunogenicity and reactivity with neutralising antibodies and has fewer
 CC side-effects when used as a drug, compared to wild type superantigen.
 XX
 XX Sequence 233 AA;

Query Match 77.3%; Score 941; DB 18; Length 233;
 Best Local Similarity 76.0%; Pred. No. 2.le-87;
 Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

Qy 1 SEKSEENEDLRKSELOGTALGNLQIYYNSKAITSEKSDAQFLTNLTFKGFETG 60
 Db 1 SEKSEENEDLRKSELOGTALGNLQIYYNEKAKTENKESHQDLQHTLTFKGFETD 60

Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMGVYTLHDNNRLT 120
 Db 61 HSWYNDLLVDLFDSDKIDVYKGGKVDLYGAYGYQCAGGTPNKTACMGVYTLHDNNRLT 120

Qy 121 EEKVPINLWIDGKQTTVPIDKVTSKKEVTYQBELDQARHYLHGKFGLYNSDSFGKGVQ 180
 Db 121 EEKVPINLWIDGKQNTVPLETATKNTKNTVQBELDQARRYLQEKYNLYNSDVFDEKGVQ 180

Qy 181 RGLVHFSSSEGSTVSYDLFDAQGOYPTDLLRIYRDNNTTISLSISLYLYTT 233
 Db 181 RGLVFTSTEPSVNYDLFGAQQOYNTLLRIYRDNKNTINSENHIDIYLYTS 233

RESULT 13
 AAB67338
 ID AAB67338 standard; peptide; 233 AA.
 XX
 AC AAB67338;

XX 23-APR-2001 (first entry)
 XX
 DE Staphylococcus aureus enterotoxin A protein.
 XX Tumour; cancer; immune; enterotoxin.
 XX Staphylococcus aureus.

XX US6180097-B1.
 XX 30-JAN-2001.
 XX 30-OCT-1998; 98US-0183437.
 XX 31-JAN-1994; 94US-0189424.
 PR 19-JUN-1995; 95US-0491746.

PR 03-OCT-1989; 89US-0416530.
 PR 17-JAN-1990; 90US-0466577.
 PR 17-JAN-1991; 91WO-US00342.
 PR 01-JUN-1992; 92US-0891718.
 PR 02-MAR-1993; 93US-0025144.
 XX (TERM/) Terman D S.
 PA Terman DS;
 XX WPI; 2001-1586657/16.
 DR
 XX Tumor cell capable of stimulating antitumor immune reactivity in vitro
 PT or in vivo comprises exogenous nucleic acids encoding a superantigen
 PT and a costimulatory molecule
 XX
 XX Disclosure; Fig 2; 16pp; English.
 XX The present invention relates to a tumour cell capable of stimulating
 CC antitumor immune reactivity in vitro or in vivo contains and
 CC expresses an exogenous nucleic acid molecule encoding a superantigen
 CC or its active fragment and an exogenous nucleic acid molecule
 CC encoding a costimulatory molecule that activates T cells in
 CC conjunction with an antigenic stimulus. The invention may be used
 CC for cancer therapy by stimulating an anticancer immune response
 CC in vivo or ex vivo.
 XX
 XX Sequence 233 AA;

Query Match 77.3%; Score 941; DB 22; Length 233;
 Best Local Similarity 76.0%; Pred. No. 2.le-87;
 Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

Qy 1 SEKSEENEDLRKSELOGTALGNLQIYYNSKAITSEKSDAQFLTNLTFKGFETG 60
 Db 1 SEKSEENEDLRKSELOGTALGNLQIYYNEKAKTENKESHQDLQHTLTFKGFETD 60

Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMGVYTLHDNNRLT 120
 Db 61 HSWYNDLLVDLFDSDKIDVYKGGKVDLYGAYGYQCAGGTPNKTACMGVYTLHDNNRLT 120

Qy 121 EEKVPINLWIDGKQTTVPIDKVTSKKEVTYQBELDQARHYLHGKFGLYNSDSFGKGVQ 180
 Db 121 EEKVPINLWIDGKQNTVPLETATKNTKNTVQBELDQARRYLQEKYNLYNSDVFDEKGVQ 180

Qy 181 RGLVHFSSSEGSTVSYDLFDAQGOYPTDLLRIYRDNNTTISLSISLYLYTT 233
 Db 181 RGLVFTSTEPSVNYDLFGAQQOYNTLLRIYRDNKNTINSENHIDIYLYTS 233

RESULT 14
 ABB76234
 ID ABB76234 standard; Protein; 233 AA.
 XX
 AC ABB76234;

XX 09-AUG-2002 (first entry)
 XX
 DE Staphylococcus aureus enterotoxin A.
 XX Enterotoxin A; SEA; superantigen; antigen; tumour; cancer;
 KW antitumour; therapy.
 XX Staphylococcus aureus.

XX Key Location/Qualifiers
 FT Misc-difference 49
 FT /note= "amino acid residue given as 'O' in the
 FT specification"
 XX US2002051765-A1.
 XX 02-MAY-2002.

```
XX PF 19-DEC-2000; 2000US-0741503.
XX PR 31-JAN-1994; 94US-0189424.
XX PR 19-JUN-1995; 95US-0491746.
XX PR 03-OCT-1989; 89US-0416530.
XX PR 17-JAN-1990; 90US-0466577.
XX PR 17-JAN-1991; 91WO-US00342.
XX PR 01-JUN-1992; 92US-0891718.
XX PR 02-MAR-1993; 93US-0025144.
XX PA (TERM/) TERMAN D S.
XX PI Terman DS;
XX DR WPI: 2002-415198/44.
XX PT Reagent for treating cancer without the need for e.g. radiotherapy,
XX PT comprises a specific V beta subset of T cells sensitized to a growing
XX PT tumor and stimulated with superantigens -
XX PS Disclosure; Fig 2; 17pp; English.
XX CC The present sequence is the protein sequence of enterotoxin A
XX CC (SEA) of Staphylococcus aureus. Similarity is shown, in several
XX CC stretches of sequence, between staphylococcal enterotoxins,
XX CC streptococcal pyrogenic exotoxins and staphylococcal exfoliative
XX CC toxins (see ABB76234-44). In the present invention, synthetic
XX CC polypeptides useful in tumour therapy and in blocking or destroying
XX CC autoreactive T and B lymphocyte populations are characterised by
XX CC substantial structural homology to staphylococcal enterotoxin A and
XX CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with
XX CC statistically significant sequence homology and similarity (Z value
XX CC of Lipman and Pearson algorithm in Monte Carlo analysis exceeding
XX CC 6) to include alignment of cysteine residues and similar hydropathy
XX CC profiles. These superantigens are used to treat solid tumours,
XX CC including their metastases, without radiation, surgery or
XX CC standard chemotherapeutic agents. A claimed method of human cancer
XX CC treatment involves contacting haematopoietic cells from a patient
XX CC with one or more superantigens ex vivo to generate stimulated cells,
XX CC selecting a specific V beta subset of cells, and reintroducing
XX CC these cells into the patient to induce an in vivo therapeutic,
XX CC tumouricidal reaction.
XX SQ Sequence 233 AA;
    Query Match 76.6%; Score 933; DB 23; Length 233;
    Best Local Similarity 75.5%; Pred. No. 1.3e-86;
    Matches 176; Conservative 21; Mismatches 36; Indels 0; Gaps 0;
QY 1 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITSSEKSADQFLNTLLFGFFTG 60
Db 1 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITENKESHDOFLXHTILFKGFFTD 60
QY 61 HPWYNDLLVDLGSTATSEYSGSSVDLYGAYCYQCAGGTPNKTCMGVTLHDNRLLT 120
Db 61 HSWYNDLLVDLDFSDKIDVRYKGGKVDLYGAYCYQCAGGTPNKTCMGVTLHDNRLLT 120
QY 121 EEKKVPINLWIDGKQTPVIDKVKTSKEVTVQELDLQARHVLHGKFGLYNSDFGKQV 180
Db 121 EEKKVPINLWLDGKQNTVPLETVTKNKNNTVQELDPQARRYLQEKYNYNSDFVFGKQV 180
QY 181 RGLIVPHSSEGSVSYDLFDAQGVPTDLLRIYRONTTISSTSLSLSLYLYTT 233
Db 181 RGLIVFTSTEPSVNYDLFCAGQVSNLTLLRIYRDNKTINSENHIDIYLYTS 233
RESULT 15
ABB79501
XX ID ABB79501 standard; Protein; 257 AA.
XX AC ABB79501;
XX DB 25 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITENKESHDOFLXHTILFKGFFTD 84
```

```
DT 23-SEP-2002 (first entry)
XX Staphylococcal enterotoxin A vaccine, periplasmic (A489270P).
XX DE Enterotoxin A; superantigen; antigen; toxin; vaccine; A489270P;
XX KW attenuation; mutant; mutein.
XX OS Staphylococcus sp.
XX OS Synthetic.
XX PH Key Location/Qualifiers
FT Peptide I..25
FT Protein /label= Signal_peptide
FT /label= Mature_protein
FT Misc-difference 72
FT /note= "wild-type Leu substituted by Arg"
FT Misc-difference 94
FT /note= "wild-type Asp substituted by Arg"
FT Misc-difference 116
FT /note= "wild-type Tyr substituted by Ala"
XX US6399332-B1.
XX 04-JUN-2002.
XX 01-SEP-1998; 98US-0144776.
XX 25-JUN-1997; 97US-0882431.
XX (USSA ) US SEC OF ARMY.
XX Ulrich RG, Olson MA, Bavari S;
XX WPI: 2002-546281/58.
XX N-PSDB; ABB84222.
XX Novel isolated and purified superantigen toxin DNA fragment which has
XX been genetically altered, useful for producing vaccine for treatment of
XX superantigen toxin-associated bacterial diseases -
XX Claim 4; Column 33-35; 46pp; English.
XX The present sequence is the protein sequence of staphylococcal
XX enterotoxin A (SEA) vaccine, periplasmic (A489270P). The vaccine
XX comprises 3 amino acid substitutions introduced into the SEA
XX sequence: L48R, Y89A and D70R. These mutations reduce the binding
XX of the toxin to major histocompatibility complex (MHC) Class II
XX and/or T cell receptors. The full-length expressed product is
XX secreted into the periplasmic space of Escherichia coli host cells,
XX and the leader peptide is recognised and cleaved by a native
XX mechanism. The vaccine is used to protect against superantigen
XX toxin infections. Superantigen attributes are absent, but the
XX superantigen is effectively recognised by the immune system and an
XX appropriate antibody response is produced. In examples from the
XX invention, attenuated superantigen toxins were shown to protect
XX animals against challenge with wild-type toxin. Methods of
XX producing and using the altered superantigen toxins as vaccines,
XX and in diagnosis and therapy, are provided. A multivalent
XX vaccine consisting of altered superantigen toxins from SEA, SEB,
XX SEC-1, TSST-1 and streptococcal SPEA is predicted to provide
XX protective immunity against the majority of bacterial superantigen
XX toxins.
XX SQ Sequence 257 AA;
    Query Match 75.9%; Score 925; DB 23; Length 257;
    Best Local Similarity 75.1%; Pred. No. 1e-85;
    Matches 175; Conservative 21; Mismatches 37; Indels 0; Gaps 0;
```

```
QY 1 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITSSEKSADQFLNTLLFGFFTG 60
Db 25 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITENKESHDOFLXHTILFKGFFTD 84
```


FT Misc-difference 2 /note= "Encoded by AG"
FT Misc-difference 18 /note= "Encoded by TTG"
FT Misc-difference 48 /note= "Wild type Leu substituted with Arg"
FT Misc-difference 70 /note= "Wild type Asp substituted with Arg"
FT Misc-difference 92 /note= "Wild type Tyr substituted with Ala"
FT Misc-difference 157 /note= "Encoded by CTT"
FT Misc-difference 180 /note= "Encoded by CAG"
XX
PN WO200009154-A1.
XX
PD 24-FEB-2000.
XX
XX 13-AUG-1998; 98WO-US16766.
XX
PR 13-AUG-1998; 98WO-US16766.
XX
PA (REED-) REED ARMY INST RES WALTER.
XX
PI Ulrich RG, Olson MA, Bavari S;
XX
XX WPI: 2000-224177/19.
DR N-PSDB; AAZ51106.
XX
XX Nucleic acid encoding superantigen toxin useful as a vaccine and for
PT diagnosis of superantigen-associated bacterial infections -
XX
XX Claim 8; Page 74-76; 118pp; English.
XX
CC The present amino acid sequence is the mutant Staphylococcal enterotoxin
CC A (SEA), a bacterial superantigen toxin (Sag), used for the
CC formulation of SEA vaccine A489270P. The coding region of this Sag toxin
CC is altered by site directed mutagenesis, that results in disruption of
CC binding of the toxin to both the MHC class II or T-cell antigen
CC receptor. This altered Sag toxin has the leader peptide cleaved by native
CC bacterial enzymatic mechanism and the first residue of the mature protein
CC is encoded by the transcriptional start site (ATG). SEA has
CC antibacterial and cytostatic activity. This sequence is useful for the
CC production of SEA vaccines and specific antibodies. This vaccine
CC overcomes the disadvantages of the chemically inactivated toxoids and is
CC designed to protect individuals against one or several related
CC staphylococcal and streptococcal toxins. It is used for the diagnosis and
CC treatment or amelioration of superantigen-associated bacterial
CC infections.
XX
SQ Sequence 233 AA;
Query Match 74.1%; Score 903; DB 21; Length 233;
Best Local Similarity 73.7%; Pred. No. 1.5e-83;
Matches 171; Conservative 21; Mismatches 40; Indels 0; Gaps 0;
QY 2 EKSEETNEKDLRKSELOQTALGNLKOIYYNSKAITSSSEKSADOFLTNTLLFGKFFTGH 61
DB 2 EKSEETNEKDLRKSEKQGTALGNLKOIYYNEKAKTENKESHDFRQUTILFGKFFTDH 61
QY 62 PNYNDLLVGLSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
DB 62 SWYNDLLVRFSKDIVDKYKGGKVDLYGAYAGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
QY 122 EKKVPINLWIDGKQTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDFSGKQVOR 181
DB 122 EKKVPINLWLDGKQNTVPLETVTKNKNVTVOELDKQARRYLQEKYNYNSDVFQKVAR 181
QY 182 GLIVFHSSSGSVSYDLDFDAQGYPDFTLLRIYRDNTTISSTLSLSLYLYTT 233
DB 182 GLIVFHTSTPSVNDLDFGAQGYQNTLLRIYRDKNKTINSENNHIDIYLYTS 233

RESULT 20
AAR45013
XX AAR45013 standard; protein; 228 AA.
XX AAR45013;
XX 08-JUN-1994 (first entry)
DT Staphylococcal enterotoxin SED.
DE
XX Staphylococcal enterotoxin SE; cancer; tumouricidal agent;
KW autoimmune disease; toxicity; Protein A; perfusion system.
XX
OS Staphylococcus aureus.
XX
PN WO9324136-A.
XX
PD 09-DEC-1993.
XX
PF 01-JUN-1993; 93WO-US05213.
XX
PR 01-JUN-1992; 92US-0891718.
XX
PA (STON/) STONE J L.
PA (TERM/) TERMAN D S.
XX
PI Stone JL, Terman DS;
XX
XX WPI: 1993-405418/50.
DR
XX Use of staphylococcal enterotoxin(s) and homologues - for
PT treating cancer in a patient or for the treatment of auto-immune
PT diseases
XX
PS Disclosure; Fig 1; 90pp; English.
XX
CC The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
CC which may be used in the methods of the invention for treating cancer
CC in a patient. These SEs, and homologues of them, can be used as
CC tumouricidal agents for treating cancers and autoimmune disease.
CC They exhibit tumouricidal activity and toxicity identical to that
CC observed for the Protein A perfusion system. They may be administered
CC by i.v. injection.
XX
SQ Sequence 228 AA;
Query Match 49.7%; Score 605; DB 14; Length 228;
Best Local Similarity 52.0%; Pred. No. 3.3e-53;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;
QY 7 INEKDLRKSELOQTALGNLKOIYYNSKAITSSSEKSADOFLTNTLLPKGFFTHPWYND 66
DB 2 VKEKELHKKSELSTALNNMKHSYADKNPIIGENKSTGDOFLENTLLYKKFFTDLINFED 61
QY 67 LLVLDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVP 126
DB 62 LLINFSKEMAOHFKSKNVDVYPIRYSINCYGGEIDRTACTYGGVTPHEGNLKERKKIP 121
QY 127 INLWIDGKQTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDFSGKQVORGLIVF 186
DB 122 INLWINGVQKESLDBKVTQDKKNVTVOELDAQARRYLQKDLKLYNNDTLGGKIQRGKIEF 181
QY 187 HSSEGSTSVSYDLDFDAQGYPDFTLLRIYRDNTTISSTLSLSLYLY 231
DB 182 DSSDGKSVSYDLDFVKGDFPEKQLRIYSDNKTLSLSTLSTLHLDIYLY 226
RESULT 21
AAB67340
ID AAB67340 standard; peptide; 228 AA.
XX
AC AAB67340;

XX 23-APR-2001 (first entry)
XX Staphylococcus aureus enterotoxin D protein.
XX Tumour; cancer; immune; enterotoxin.
XX Staphylococcus aureus.
XX US6180097-B1.
XX 30-JAN-2001.
XX 30-OCT-1998; 98US-0183437.
XX 31-JAN-1994; 94US-0189424.
XX 19-JUN-1995; 95US-0491746.
XX 03-OCT-1989; 89US-0416530.
XX 17-JAN-1990; 90US-0466577.
XX 17-JAN-1991; 91WO-US00342.
XX 01-JUN-1992; 92US-0891718.
XX 02-MAR-1993; 93US-0025144.
XX (TERM/) TERMAN D S.
XX Terman DS;
XX WPI; 2001-158657/16.
XX Tumor cell capable of stimulating antitumor immune reactivity in vitro
PT or in vivo comprises exogenous nucleic acids encoding a superantigen
PT and a costimulatory molecule -
XX Disclosure; Fig 2; 16pp; English.
XX The present invention relates to a tumour cell capable of stimulating
CC antitumor immune reactivity in vitro or in vivo contains and
CC expresses an exogenous nucleic acid molecule encoding a superantigen
CC or its active fragment and an exogenous nucleic acid molecule
CC encoding a costimulatory molecule that activates T cells in
CC conjunction with an antigenic stimulus. The invention may be used
CC for cancer therapy by stimulating an anticancer immune response
CC in vivo or ex vivo.
XX Sequence 228 AA;
Query Match 49.7%; Score 605; DB 22; Length 228;
Best Local Similarity 52.0%; Pred. No. 3.3e-53;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;
Qy 7 INEKDLRKSELOGTALGNLKOIYYVNSKAITSSPKSADQFLTNTLLPKGFTGHPWYND 66
Db 2 VKERELHKSELSSTALNMKHSYADKNPIIGENKSTGDFLENTLLYKFFTDLINEF 61
Qy 67 LLVDLGSTAATSEYEGSSVDLYGAYGYQCAGTTPNKACTMGYVTLHDNNRLTEKKVP 126
Db 62 LLINPNSKEMAQHFKSNVDVPIRYSINCYGGEIDRTACTYGGVTPHEGNKLERKKIP 121
Qy 127 INLWIDGKQTTVPIDKVKTSKKEVTQVQLDQARHYLHGKFGLYNSDFGKVGRLIVF 186
Db 122 INLWINGVQKESLDKVTQDKNNVTQVQLDQARHYLHGKFGLYNSDFGKVGRLIVF 186
Qy 187 HSEGSTSYDLFDAQGYPTDTPLLRIYRDNTTISSTLSISLYL 231
Db 182 DSSDGSKSYDLFDVKGDFPEKQLRIYSDNKTLSLHIDILY 226
RESULT 22
ABB76236
ID ABB76236 standard; Protein; 228 AA.
XX
AC ABB76236;
XX

DT 09-AUG-2002 (first entry)
XX Staphylococcus aureus enterotoxin D.
DE Enterotoxin D; SED; superantigen; antigen; tumour; cancer;
XX antitumour; therapy.
KW Staphylococcus aureus.
XX US2002051765-A1.
XX 02-MAY-2002.
XX 19-DEC-2000; 2000US-0741503.
XX 31-JAN-1994; 94US-0189424.
XX 19-JUN-1995; 95US-0491746.
XX 03-OCT-1989; 89US-0416530.
XX 17-JAN-1990; 90US-0466577.
XX 17-JAN-1991; 91WO-US00342.
XX 01-JUN-1992; 92US-0891718.
XX 02-MAR-1993; 93US-0025144.
XX (TERM/) TERMAN D S.
PA Terman DS;
XX WPI; 2002-415198/44.
XX Reagent for treating cancer without the need for e.g. radiotherapy,
PT comprises a specific V beta subset of T cells sensitized to a growing
PT tumor and stimulated with superantigens -
XX Disclosure; Fig 2; 17pp; English.
XX The present sequence is the protein sequence of enterotoxin D
CC (SED) of Staphylococcus aureus. Similarity is shown, in several
CC stretches of sequence, between staphylococcal enterotoxins,
CC streptococcal pyrogenic exotoxins and staphylococcal exfoliative
CC toxins (see ABB76234-44). In the present invention, synthetic
CC polypeptides useful in tumour therapy and in blocking or destroying
CC autoreactive T and B lymphocyte populations are characterized by
CC substantial structural homology to staphylococcal enterotoxin A and
CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with
CC statistically significant sequence homology and similarity (Z value
CC of Lipman and Pearson algorithm in Monte Carlo analysis exceeding
CC 6) to include alignment of cysteine residues and similar hydrophathy
CC profiles. These superantigens are used to treat solid tumours,
CC including their metastases, without radiation, surgery or
CC standard chemotherapeutic agents. A claimed method of human cancer
CC treatment involves contacting haematopoietic cells from a patient
CC with one or more superantigens ex vivo to generate stimulated cells,
CC selecting a specific V beta subset of cells, and reintroducing
CC these cells into the patient to induce an in vivo therapeutic,
CC tumouricidal reaction.
XX Sequence 228 AA;
Query Match 49.7%; Score 605; DB 23; Length 228;
Best Local Similarity 52.0%; Pred. No. 3.3e-53;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;
Qy 7 INEKDLRKSELOGTALGNLKOIYYVNSKAITSSPKSADQFLTNTLLPKGFTGHPWYND 66
Db 2 VKERELHKSELSSTALNMKHSYADKNPIIGENKSTGDFLENTLLYKFFTDLINEF 61
Qy 67 LLVDLGSTAATSEYEGSSVDLYGAYGYQCAGTTPNKACTMGYVTLHDNNRLTEKKVP 126
Db 62 LLINPNSKEMAQHFKSNVDVPIRYSINCYGGEIDRTACTYGGVTPHEGNKLERKKIP 121
Qy 127 INLWIDGKQTTVPIDKVKTSKKEVTQVQLDQARHYLHGKFGLYNSDFGKVGRLIVF 186
Db 122 INLWINGVQKESLDKVTQDKNNVTQVQLDQARHYLHGKFGLYNSDFGKVGRLIVF 186

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI: 2002-352536/38.
DR N-PSDB; ABN69988.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 3920; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 259 AA;
Query Match 23.9%; Score 291.5; DB 23; Length 259;
Best Local Similarity 31.3%; Pred. No. 3.3e-21;
Matches 72; Conservative 42; Mismatches 91; Indels 25; Gaps 8;
QY 17 ELQGTALGNLKIY-YNNSKATTSSEKSADQFLTNLLFKGFTGHPWVN---DLIVDLG 72
DB 34 ENSSVGVINLRNLSTYDPTVEVKGINSGPP-SCSLFYKNT----PYGNSIELKVELN 88
QY 73 STAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWID 132
DB 89 SVEKANFESGKRVDIFTLEYSPPCNSIKNS---YGGITLSDGNRI-DKNKIPVNIIFD 144
QY 133 G---KQTTVPIDKVTSKKEVTVOELDLQARHYLHGKGLG-----NSDSFGGKVQ 180
DB 145 GVQQKYSYTDISTGTDKKEVTIQELDKRSYLLQKHENIYFGVDKDFGRSSRQSGFE 204
QY 181 RGLIVFHSSEGSTVSDYDLFDAQGQYPTDLLRIYRDNNTTSSLSLSLYL 230
DB 205 EGNIIIFHLSGERISYNLFDTGCHGDRSMLKKYSNKTAYSQDLHIDIYL 254

RESULT 26
AA06254
ID AAY06254 standard; Protein; 239 AA.
XX
AC AAY06254;
XX
DT 23-AUG-1999 (first entry)
XX
DE Staphylococcal group C enterotoxin SEC3-FRI909.
KW Enterotoxin; SEC3-FRI909; toxin; disulfide loop;
KW protein engineering.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 93..110
XX
PN WO9927889-A2.
XX
PD 10-JUN-1999.
XX
PF 01-DEC-1998; 98WO-US25107.
XX
PR 02-DEC-1997; 97US-0067357.
XX
PA (IDAH-) IDAHO RES FOUND INC.
XX
PI Bohach GI;
XX
DR WPI: 1999-358008/30.
XX
XX Non-toxic modified staphylococcal enterotoxins
PS Disclosure; Page 17; 25pp; English.
XX
XX This protein represents the Staphylococcus aureus type C
CC enterotoxin SEC3-FRI909. The invention relates to pyrogenic toxins,
CC such as staphylococcal enterotoxins, modified in the disulfide loop
CC region. Typically, the modification involves deletions within the
CC disulfide loop region of SEC (see AAY06261). The modified toxins
CC retain useful biological properties, such as the ability to induce
CC cytokine production, but have substantially reduced toxicity
CC compared to the corresponding unmodified native toxin. Emetic
CC response inducing activity and fever inducing activity are
CC typically decreased by at least about 100-fold, while LD50 (in Dutch
CC Belted rabbits) is at least 100-fold higher than the native toxin.
XX
SQ Sequence 239 AA;
Query Match 23.6%; Score 287.5; DB 20; Length 239;
Best Local Similarity 33.3%; Pred. No. 7.5e-21;
Matches 78; Conservative 45; Mismatches 92; Indels 19; Gaps 9;
QY 11 DLRKSELOQTALGNLKIYNNKAITSSSE-KSADQFLTNLLFKGFTGHPWYNDLLV 69
DB 10 DLHKSEETGT-MGNMK--YLYDDHYVSNATKVSVDKFLAHLDIYNINDKLNNDKYKT 56
QY 70 DLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGGVTLHDNNRLTEE 122
DB 67 ELLNEDLANKYKDEVDVYGSNNYVNCYFSSKDNVSKVTSKGTCTMGYGTKEGHNFDNG 126
QY 123 --KKYPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKGLGSDSFGGKVQ 180
DB 127 NLQNLIRYV-ENKRNITISFE-VQTDKKSVTAAQELDIKARNFLINKNLNLYEFS--SPYE 182
QY 181 RGLIVFHSSEGSTVSDYDLFDAQGQYPT--TLIRIYRDNNTTSSLSLSLYLT 232
DB 183 TCYIKFIESNGNTFWYDMMPAPGDKFDQSKYLMYKDNKMWDSKSVKIEVHLTT 236
RESULT 27
AA06253

DT 23-AUG-1999 (first entry)
XX Staphylococcal group C enterotoxin SEC-MNCopeland.
DE Enterotoxin; SEC-MNCopeland; toxin; disulfide loop;
XX protein engineering.
KW Staphylococcus aureus.
XX Key Location/Qualifiers
XX Disulfide-bond 93..110
XX WO9927889-A2.
XX 10-JUN-1999.
XX 01-DEC-1998; 98WO-US25107.
XX 02-DEC-1997; 97US-0067357.
XX (IDAH-) IDAHO RES FOUND INC.
XX Bohach GI;
XX WPI; 1999-358008/30.
XX Non-toxic modified staphylococcal enterotoxins
XX Disclosure; Page 17; 25pp; English.

XX AAY06252;
XX 23-AUG-1999 (first entry)
XX Staphylococcal group C enterotoxin SEC2.
XX Enterotoxin; SEC2; toxin; disulfide loop; protein engineering.
XX Staphylococcus aureus.
XX Key Location/Qualifiers
XX Disulfide-bond 93..110
XX WO9927889-A2.
XX 10-JUN-1999.
XX 01-DEC-1998; 98WO-US25107.
XX 02-DEC-1997; 97US-0067357.
XX (IDAH-) IDAHO RES FOUND INC.
XX Bohach GI;
XX WPI; 1999-358008/30.
XX Non-toxic modified staphylococcal enterotoxins
XX Disclosure; Page 17; 25pp; English.

CC This protein represents the Staphylococcus aureus type C
CC enterotoxin SEC-MNCopeland. The invention relates to pyrogenic
CC toxins, such as staphylococcal enterotoxins, modified in the
CC disulfide loop region. Typically, the modification involves
CC deletions within the disulfide loop region of SEC (see AAY06261).
CC The modified toxins retain useful biological properties, such as
CC the ability to induce cytokine production, but have substantially
CC reduced toxicity compared to the corresponding unmodified native
CC toxin. Emetic response inducing activity and fever inducing
CC activity are typically decreased by at least about 100-fold, while
CC LD50 (in Dutch Belted rabbits) is at least 100-fold higher than
CC the native toxin.

CC This protein represents the Staphylococcus aureus type C
CC enterotoxin SEC2. The invention relates to pyrogenic toxins, such
CC as staphylococcal enterotoxins, modified in the disulfide loop
CC region. Typically, the modification involves deletions within the
CC disulfide loop region of SEC (see AAY06261). The modified toxins
CC retain useful biological properties, such as the ability to induce
CC cytokine production, but have substantially reduced toxicity
CC compared to the corresponding unmodified native toxin. Emetic
CC response inducing activity and fever inducing activity are
CC typically decreased by at least about 100-fold, while LD50 (in Dutch
CC Belted rabbits) is at least 100-fold higher than the native toxin.

XX Sequence 239 AA;
Query Match 23.0%; Score 280.5; DB 20; Length 239;
Best Local Similarity 31.3%; Pred. No. 3.8e-20;
Matches 76; Conservative 46; Mismatches 102; Indels 19; Gaps 9;
Qy 2 EKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSE-KSADQFLNTLLFKGFTG 60
Db 1 ESQPDPTDELHKSEFTGT-MGNMK--YLYDDHYVSATKVSVDKFLAHLDIYNSDKK 57
Qy 61 HPWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACHYGGVTL 113
Db 58 LKNYDKVKTTELLNEDLAKKYKDEVDVYGSNYVNCYFESSKDNVKGKVTGCTMYGGITK 117
Qy 114 HDNNRLTEE--KKVPINLWIDGKQTTPIDKVKTSKEVTVQELDLQARHVLHGKFGLYN 171
Db 118 HEGNHFNGNQLNQLIRVY-ENKRTISFE-VQTDKKSVTQAQELDIKARNFINKKNLYE 175
Qy 172 SDSFGGKVQORGLIVFHSSEGSTVSVDLFAQGVDP--TLRLIYRDNTTISSTLSISLY 229
Db 176 FNS--SPVETGYIKFIENNGNTFOYDMMPAPGDKFQDSKYLMMYNDKNTVDSKRVKIEVH 233
Qy 230 LYT 232
Db 234 LTT 236

RESULT 30
ID AAY06252 standard; Protein; 239 AA.

RESULT 31
AAR45016

KW	Enterotoxin C3; SEC3; superantigen; antigen; tumour; cancer;
KX	antitumour; therapy.
OS	Staphylococcus aureus.
FH	Key Location/Qualifiers
FT	Misc-difference 208 /note= "given as 'o' in the specification"
FN	US2002051765-A1.
PB	02-MAY-2002.
PP	19-DEC-2000; 2000US-0741503.
PR	31-JAN-1994; 94US-0189424.
PR	19-JUN-1995; 95US-0491746.
PR	03-OCT-1989; 89US-0416530.
PR	17-JAN-1990; 90US-0466577.
PR	17-JAN-1991; 91WO-US00342.
PR	01-JUN-1992; 92US-0891718.
PR	02-MAR-1993; 93US-0025144.
PA	(TERM/) TERMAN D S.
PI	Terman DS;
XI	WPI; 2002-415198/44.
DR	Reagent for treating cancer without the need for e.g. radiotherapy,
PT	comprises a specific v beta subset of T cells sensitized to a growing
PT	tumor and stimulated with superantigens -
XX	Disclosure; Fig 2; 17pp; English.
PS	The present sequence is the protein sequence of enterotoxin C3
CC	(SEC3) of Staphylococcus aureus. Similarity is shown, in several
CC	stretches of sequence, between staphylococcal enterotoxins,
CC	streptococcal pyrogenic exotoxins and staphylococcal exfoliative
CC	toxins (see AB576234-44). In the present invention, synthetic
CC	polypeptides useful in tumor therapy and in blocking or destroying
CC	autoreactive T and B lymphocyte populations are characterized by
CC	substantial structural homology to staphylococcal enterotoxin A and
CC	enterotoxin B, and to streptococcal pyrogenic exotoxins, with
CC	statistically significant sequence homology and similarity (% value
CC	of Lipman and Pearson algorithm in Monte Carlo analysis exceeding
CC	6) to include alignment of cysteine residues and similar hydropathy
CC	profiles. These superantigens are used to treat solid tumors,
CC	including their metastases, without radiation, surgery or
CC	standard chemotherapeutic agents. A claimed method of human cancer
CC	treatment involves contacting haematopoietic cells from a patient
CC	with one or more superantigens ex vivo to generate stimulated cells,
CC	selecting a specific v beta subset of cells, and reintroducing
CC	these cells into the patient to induce an in vivo therapeutic,
CC	tumoricidal reaction.
XX	
SQ	Sequence 238 AA;
Query Match	22.6%; Score 275.5; DB 23; Length 238;
Best Local Similarity	32.1%; Pred. No. 1.2e-19;
Matches 75; Conservative	44; Mismatches 96; Indels 19; Gaps
QY	11 DLRRKSELOQTALGNKOIIYYNKAITSS-E-KSADEFLTNTLLFKGFFTGHPWNLDLV 69 : : : : : : :
Db	9 ELHKSSFECT-GMNMK--VLYDDHVVSATKMVSVDKFLAHDLTINISDKKLKNYDKVT 65 : : : : : : : : : : : :
QY	70 DLGASTATSEYGSSDLYGYTYQC-----AGTTPNKATCMYGGVTLHDNNRLTEE 122 : : : : : : : : : : : :
Db	66 ELLNEDLAKKYKDEVVDVGNSYVVNCYFSSKDVKCVGTGCKTKMGGITKEGHNFDPG 125 : : : : : : : : : :
QY	123 --KKVPINLWDGQTTVPIDIKYKTTSKKEVTVQELDLQAHHYLHGFGYSDSFGGKVQ 180 : : : : : : : : : :

QY 61 HPWYNDLLVDLGSATSEYEGSSVDLYGAYYGYQC-AGGTPNKTACMYGGVTLHDNNRL 119
DB 82 GPNYDKLTKELKNQEMATLFDKKNVDIYGVEYHLCYLSENAERSACIYGGVTNHEGNHL 141
QY 120 TEEKKVPINLWDIGQTTVPIDKVTSKKEVTQELDLQARHYLHGKFGLYNSDSFGGKV 179
DB 142 ELPKIVKVSIDGQ-SLSFD-IETNKKWTAQELDYKVRKYLTNDNKQLYTNGP--SKY 197
QY 180 QRLIVFHSSEGSTVSYDLFD----AQGYQPDTLRLIYRONTTISSTLSISLYLT 232
DB 198 ETGYIKFIPKNKESEWDFPEPEFTQSKY----LMIYKNETLDSNTSQIEVYLT 250

RESULT 37
AAR13209
ID AAR13209 standard; Protein; 221 AA.
XX
AC AAR13209;
XX
DT 15-OCT-1991 (first entry)
XX
DE Streptococcal pyrogenic enterotoxin A.
XX
KW SPE A; cancer treatment; pyrogen; tumouricide; scarlet fever.
XX
OS Streptococcus NY-5 strain.
XX
PN W09110680-A.
XX
PD 25-JUL-1991.
XX
PF 17-JAN-1991; 91WO-US00342.
XX
PR 17-JAN-1990; 90US-0466577.
XX
PA (TERM/) TERMAN D S.
XX
PI Terman DS;
XX
DR WPI; 1991-237984/32.
XX
PT Treating cancer with enterotoxin from Staphylococcus aureus -
PT administered by IV injection, having same tumoricidal activity
PT as Staphylococcal protein A without potential toxic reactions
XX
PS Disclosure; Fig 1: 74pp; English.
XX
CC SPE A can be used for tumoricidal treatment, esp. with a haemolysin.
CC Synthetic polypeptides having structural homology to Streptococcal
CC pyrogenic exotoxins are claimed, provided the homology includes
CC statistically significant sequence homology, alignment of Cysteine
CC residues and similar hydrophathy profiles.
XX
SQ See AAR13203-R13211.
XX
SQ Sequence 221 AA;
Query Match 22.2%; Score 270.5; DB 12; Length 221;
Best Local Similarity 33.8%; Pred. No. 3.6e-19;
Matches 76; Conservative 42; Mismatches 88; Indels 19; Gaps 10;

QY 16 SELOGTAL-GNLKQIYY-YNSKAIT-SSEKSDAQFLTNTLLFKGFFTHPWNVDLLVDLG 72
DB 7 SOLHRSSLVKNLQNIYFLYEGDPVTHENVKSVQDQLSHDLIYN---VSGPNYDKLTELK 63
QY 73 STAATSEYEGSSVDLYGAYYGYQC-AGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWI 131
DB 64 NQEMATLFDKKNVDIYGVEYHLCYLSENAERSACIYGGVTNHEGNHLEIPKKIYVKVSI 123
QY 132 DGKQTTVPIDKVKTSKEVTQELDLQARHYLHGKFGLYNSDSFGGKVORGLIVFHSSEG 191
DB 124 DGIQ-SLSFD-IETNKKWTAQELDYKVRKYLTNDNKQLYTNGP--SKYETGYIKFIPKNK 179
QY 192 STVSYDLFD----AQGYQPDTLRLIYRONTTISSTLSISLYLT 232

DB 180 ESFWEDLPEPEFTQSKY----LMIYKNETLDSNTSQIEVYLT 220

RESULT 38
AAR45017
ID AAR45017 standard; protein; 221 AA.
XX
AC AAR45017;
XX
DT 08-JUN-1994 (first entry)
XX
DE Staphylococcal enterotoxin SPE A.
XX
KW Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
KW autoimmune disease; toxicity; Protein A; perfusion system.
XX
OS Staphylococcus aureus.
XX
PN W09324136-A.
XX
PD 09-DEC-1993.
XX
PF 01-JUN-1993; 93WO-US05213.
XX
PR 01-JUN-1992; 92US-0891718.
XX
PA (STON/) STONE J L.
PA (TERM/) TERMAN D S.
XX
PI Stone JL, Terman DS;
XX
DR WPI; 1993-405418/50.
XX
PT Use of staphylococcal enterotoxin(s) and homologues - for
PT treating cancer in a patient or for the treatment of auto-immune
PT diseases
XX
PS Disclosure; Fig 1: 90pp; English.
XX
CC The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
CC which may be used in the methods of the invention for treating cancer
CC in a patient. These SEs, and homologues of them, can be used as
CC tumoricidal agents for treating cancers and autoimmune disease.
CC They exhibit tumoricidal activity and toxicity identical to that
CC observed for the protein A perfusion system. They may be administered
CC by i.v. injection.
XX
SQ Sequence 221 AA;
Query Match 22.2%; Score 270.5; DB 14; Length 221;
Best Local Similarity 33.8%; Pred. No. 3.6e-19;
Matches 76; Conservative 42; Mismatches 88; Indels 19; Gaps 10;

QY 16 SELOGTAL-GNLKQIYY-YNSKAIT-SSEKSDAQFLTNTLLFKGFFTHPWNVDLLVDLG 72
DB 7 SOLHRSSLVKNLQNIYFLYEGDPVTHENVKSVQDQLSHDLIYN---VSGPNYDKLTELK 63
QY 73 STAATSEYEGSSVDLYGAYYGYQC-AGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWI 131
DB 64 NQEMATLFDKKNVDIYGVEYHLCYLSENAERSACIYGGVTNHEGNHLEIPKKIYVKVSI 123
QY 132 DGKQTTVPIDKVKTSKEVTQELDLQARHYLHGKFGLYNSDSFGGKVORGLIVFHSSEG 191
DB 124 DGIQ-SLSFD-IETNKKWTAQELDYKVRKYLTNDNKQLYTNGP--SKYETGYIKFIPKNK 179
QY 192 STVSYDLFD----AQGYQPDTLRLIYRONTTISSTLSISLYLT 232
DB 180 ESFWEDLPEPEFTQSKY----LMIYKNETLDSNTSQIEVYLT 220

RESULT 39
ABB76240

Query Match 22.2%; Score 270.5; DB 18; Length 251;
Best Local Similarity 32.5%; Pred. No. 4.3e-19;
Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SBEINEKDLRKSELOGTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLNTLLFKGFFTG 60
Db 25 SBEVFAQQDPDSQLHRSSLVNQLNIYFLYEGDPVTHENVKSVQDLLSHHIIYN--VS 81
QY 61 HPWYNDDLVLGTAATSEYEGSSVDLYGAYYQOC-AGGTPNKTKACMYGGVTLHDNNRL 119
Db 82 GPNYDKLKTLELNQEMATLFLKKNVDIYGVEYHLCYLCEAERSACIYGGVTNHEGHN 141
QY 120 TEEKKVPINLWIDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKV 179
Db 142 EIPKKIIVKVSIDGQ-SLSFD-IETNKKMVTQAELDYKVRKYLTDNKQLYTNGP--SKY 197
QY 180 ORGLIIVHSESGTSVSYDLFD----AOGQYPTDLLRIYRDNNTTISSTLSISLYLT 232
Db 198 ETGYIKFIPKNKESFWDFPEFTQSKY----LMIYKDNETLDSNTSQIEVYLTT 250

RESULT 41
AAW12151
ID AAW12151 standard; Protein; 251 AA.
AC AAW12151;
DT 04-NOV-1997 (first entry)
XX Streptococcus pyogenes Streptococcal toxin A mutant Cys87Ser.
DE Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
KW vaccine; protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
KW fever; hypotension; group A streptococcal infection; myositis;
KW fascitis; liver damage; T cell; lymphoma; ovarian; uterine.
XX Streptococcus pyogenes.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT /label= sig_peptide
FT Peptide 31..251
FT /label= mat_peptide
FT Misc-difference 117
FT /note= "wild type Cys replaced by Ser"
XX
PN WO9640930-Al.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US10252.
XX
PR 07-JUN-1995; 95US-0480261.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
XX WPI; 1997-099936/09.
XX
PT Mutant SPE-A toxin with at least one amino acid change is
PT substantially non-lethal - used in vaccine composition for
PT treatment of cancer and streptococcal toxic shock syndrome etc.
XX
PS Example 4; Page -: 102pp; English.
XX
CC The present sequence is a non-lethal Streptococcus pyogenes
CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce
CC vaccines to protect animals against wild type SPE-A and to treat
CC cancer and streptococcal toxic shock syndrome (STSS). The mutant

CC SPE-A causes neutralising antibodies (Ab) to be produced, which
CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,
CC group A streptococcal infection, myositis, fascitis and liver
CC damage. The neutralising Ab is preferably administered in
CC conjunction with antibiotic therapy. The mutant SPE-A is
CC especially useful for treating T cell lymphomas, and ovarian and
CC uterine cancer. It is thought that mutant SPE-A can be selectively
CC toxic to T cell lymphoma cells.
CC N.B. Sequence not given in the specification, but constructed
CC using the wild type SPE-A sequence given on pages 77-79.
XX
SQ Sequence 251 AA;

Query Match 22.2%; Score 270.5; DB 18; Length 251;
Best Local Similarity 32.6%; Pred. No. 4.3e-19;
Matches 78; Conservative 44; Mismatches 94; Indels 23; Gaps 11;

QY 4 SBEINEKDLRKSELOGTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLNTLLFKGFFTG 60
Db 25 SBEVFAQQDPDSQLHRSSLVNQLNIYFLYEGDPVTHENVKSVQDLLSHHIIYN--VS 81
QY 61 HPWYNDDLVLGTAATSEYEGSSVDLYGAYY---GYOCAGGTPNKTKACMYGGVTLHDNN 117
Db 82 GPNYDKLKTLELNQEMATLFLKKNVDIYGVEYHLSYLCENA--ERSACIYGGVTNHEGN 139
QY 118 RLTEKKVPINLWIDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGG 177
Db 140 HLEIPKKIIVKVSIDGQ-SLSFD-IETNKKMVTQAELDYKVRKYLTDNKQLYTNGP--S 195
QY 178 KVQGLIIVHSESGTSVSYDLFD----AOGQYPTDLLRIYRDNNTTISSTLSISLYLT 232
Db 196 KYETGYIKFIPKNKESFWDFPEFTQSKY----LMIYKDNETLDSNTSQIEVYLTT 250

RESULT 42
AAW12154
ID AAW12154 standard; Protein; 251 AA.
XX
AC AAW12154;
XX
DT 04-NOV-1997 (first entry)
XX Streptococcus pyogenes Streptococcal toxin A mutant Ser195Ala.
DE Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
KW vaccine; protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
KW fever; hypotension; group A streptococcal infection; myositis;
KW fascitis; liver damage; T cell; lymphoma; ovarian; uterine.
XX Streptococcus pyogenes.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT /label= sig_peptide
FT Peptide 31..251
FT /label= mat_peptide
FT Misc-difference 225
FT /note= "wild type Ser replaced by Ala"
XX
PN WO9640930-Al.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US10252.
XX
PR 07-JUN-1995; 95US-0480261.
XX (MINU) UNIV MINNESOTA.
PI Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
XX

DR WPI; 1997-099936/09.
XX
XX Mutant SPE-A toxin with at least one amino acid change is
PT substantially non-lethal - used in vaccine composition for
PT treatment of cancer and streptococcal toxic shock syndrome etc.
XX
XX
XX Example 4; Page -: 102pp; English.
XX
XX The present sequence is a non-lethal Streptococcus pyogenes
CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce
CC vaccines to protect animals against wild type SPE-A and to treat
CC cancer and streptococcal toxic shock syndrome (STSS). The mutant
CC SPE-A causes neutralising antibodies (Ab) to be produced, which
CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,
CC group A streptococcal infection, myositis, fascitis and liver
CC damage. The neutralising Ab is preferably administered in
CC conjunction with antibiotic therapy. The mutant SPE-A is
CC especially useful for treating T cell lymphomas, and ovarian and
CC uterine cancer. It is thought that mutant SPE-A can be selectively
CC toxic to T cell lymphoma cells.
CC N.B. Sequence not given in the specification, but constructed
CC using the wild type SPE-A sequence given on pages 77-79.
XX
XX Sequence 251 AA;
SQ
Query Match 22.0%; Score 268.5; DB 18; Length 251;
Best Local Similarity 32.5%; Pred. No. 6.9e-19;
Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;
QY 4 SEEINEKLRKSELGQAL-GNLKOIYY-YNSKAIT-SSEKSADQFLTNTLLKGFFTG 60
Db 25 SOEFAAQDPDPSQLHRSRLNQLNIYFLYEGDPVTHENVKSVQDLSHLLIYN---VS 81
QY 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOC-AGGTPNKACMYGGVTLHDNNRL 119
Db 82 GPNYDKLTKELKNQEMATLFDKNDVIYGVVEYHLCYLCEAERSACIYGGVTNHEGNHL 141
QY 120 TEEKKVPINLWDGKQTTVPIDKVKTSKKEVTQVELDQARHYLHGKFLGYNDSFSGGKV 179
Db 142 EIPKKIVKVSIDGIO-SLSFD-IETNKKMVTQAEIDYKVKYLDNKOLYTNGP--SKY 197
QY 180 QRLIVFHSSEGSTVSYDLFD----AQQYPTLLRIYRDNNTTISSTLSISLYLT 232
Db 198 ETGYIKFIPKNKESFWDFPEPEFTQAKY----LMIYKDNELTSDNTSQIEVILTT 250
RESULT 43
AAW12148
ID AAW12148 standard; Protein; 251 AA.
XX
XX AAW12148;
XX
XX 04-NOV-1997 (first entry)
XX Streptococcus pyogenes Streptococcal toxin A mutant N20D/K157E.
XX
XX Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
KW vaccine; protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
KW fever; hypotension; group A streptococcal infection; myositis;
KW fascitis; liver damage; T cell; lymphoma; ovarian; uterine.
XX
XX Streptococcus pyogenes.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Peptide 1..30
FT /label= sig_peptide
FT Peptide 31..251
FT /label= mat_peptide
FT Misc-difference 50
FT /note= "wild type Asn replaced by Asp"
FT Misc-difference 187

FT /note= "wild type Lys replaced by Glu"
XX
XX WO9640930-A1.
XX 19-DEC-1996.
PD
XX
XX 07-JUN-1996; 96WO-US10252.
XX
XX 07-JUN-1995; 95US-0480261.
XX (MINU) UNIV MINNESOTA.
PA
XX Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
PI WPI; 1997-099936/09.
DR
XX Mutant SPE-A toxin with at least one amino acid change is
PT substantially non-lethal - used in vaccine composition for
PT treatment of cancer and streptococcal toxic shock syndrome etc.
XX
XX Claim 5; Page -: 102pp; English.
PS
XX The present sequence is a non-lethal Streptococcus pyogenes
CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce
CC vaccines to protect animals against wild type SPE-A and to treat
CC cancer and streptococcal toxic shock syndrome (STSS). The mutant
CC SPE-A causes neutralising antibodies (Ab) to be produced, which
CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,
CC group A streptococcal infection, myositis, fascitis and liver
CC damage. The neutralising Ab is preferably administered in
CC conjunction with antibiotic therapy. The mutant SPE-A is
CC especially useful for treating T cell lymphomas, and ovarian and
CC uterine cancer. It is thought that mutant SPE-A can be selectively
CC toxic to T cell lymphoma cells.
CC N.B. Sequence not given in the specification, but constructed
CC using the wild type SPE-A sequence given on pages 77-79.
XX
XX Sequence 251 AA;
SQ
Query Match 22.0%; Score 268.5; DB 18; Length 251;
Best Local Similarity 32.5%; Pred. No. 6.9e-19;
Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;
QY 4 SEEINEKLRKSELGQAL-GNLKOIYY-YNSKAIT-SSEKSADQFLTNTLLKGFFTG 60
Db 25 SOEFAAQDPDPSQLHRSRLNQLNIYFLYEGDPVTHENVKSVQDLSHLLIYN---VS 81
QY 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOC-AGGTPNKACMYGGVTLHDNNRL 119
Db 82 GPNYDKLTKELKNQEMATLFDKNDVIYGVVEYHLCYLCEAERSACIYGGVTNHEGNHL 141
QY 120 TEEKKVPINLWDGKQTTVPIDKVKTSKKEVTQVELDQARHYLHGKFLGYNDSFSGGKV 179
Db 142 EIPKKIVKVSIDGIO-SLSFD-IETNKKMVTQAEIDYKVKYLDNKOLYTNGP--SKY 197
QY 180 QRLIVFHSSEGSTVSYDLFD----AQQYPTLLRIYRDNNTTISSTLSISLYLT 232
Db 198 ETGYIKFIPKNKESFWDFPEPEFTQSKY----LMIYKDNELTSDNTSQIEVILTT 250
RESULT 44
AAW12147
ID AAW12147 standard; Protein; 251 AA.
XX
XX AAW12147;
XX
XX 04-NOV-1997 (first entry)
XX Streptococcus pyogenes Streptococcal toxin A mutant Asn20Asp.
XX Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
KW vaccine; protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;

KW fever; hypotension; group A streptococcal infection; myositis;
 KW fasciitis; liver damage; T cell; lymphoma; ovarian; uterine.
 XX Streptococcus pyogenes.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Peptide 1..30
 FT /label= sig_peptide
 FT Peptide 31..251
 FT /label= mat_peptide
 FT Misc-difference 50
 FT /note= "wild type Asn replaced by Asp"

XX WO9640930-A1.
 XX 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US10252.
 XX 07-JUN-1995; 95US-0480261.
 XX (MINU) UNIV MINNESOTA.
 XX Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
 XX WPI; 1997-099936/09.
 XX Mutant SPE-A toxin with at least one amino acid change is
 PT substantially non-lethal - used in vaccine composition for
 PT treatment of cancer and streptococcal toxic shock syndrome etc.
 XX Claim 5; Page -: 102pp; English.

XX The present sequence is a non-lethal Streptococcus pyogenes
 CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce
 CC vaccines to protect animals against wild type SPE-A and to treat
 CC cancer and streptococcal toxic shock syndrome (STSS). The mutant
 CC SPE-A causes neutralising antibodies (Ab) to be produced, which
 CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,
 CC group A streptococcal infection, myositis, fasciitis and liver
 CC damage. The neutralising Ab is preferably administered in
 CC conjunction with antibiotic therapy. The mutant SPE-A is
 CC especially useful for treating T cell lymphomas, and ovarian and
 CC uterine cancer. It is thought that mutant SPE-A can be selectively
 CC toxic to T cell lymphoma cells.
 CC N.B. Sequence not given in the specification, but constructed
 CC using the wild type SPE-A sequence given on pages 77-79.

XX Sequence 251 AA;
 Query Match 22.0%; Score 268.5; DB 18; Length 251;
 Best Local Similarity 32.5%; Pred. No. 6.9e-19;
 Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SEINEKDLRKXSELOGTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLNTLLFGFFTG 60
 Db 25 SOEVAQQDPDSQLHRSSLVKNLQDIPLYEGDPVTHENVKSVQQLSHLLIYN--VS 81
 QY 61 HPWYNLLVDLGLSTAAATSEYEGSSVDLYGAYYGYOC-AGGTPNKTKACYGGVTLHDNNRL 119
 Db 82 GPNYDKLKTENKQEMATLFDKKNVDIYGVYHYLCYLCENASACIYGGVTNHEGNHL 141
 QY 120 TEEKKVPINLWDGKQTTVPIDKVKTSKEVTQVLELDQARHLHGKFLGNSDFGGKV 179
 Db 142 ETPKILVVKVSDIGIQ-SUSFD-IETNKKMVAQLELDYKRVKLYLDNKKOLYTNGP--SKY 197
 QY 180 QRGLIVFHSSEGSTVSYDLFD----AQGOYPTDLRLRIYRDNTTISSTLSISLYLT 232
 Db 198 ETGYIKFIPKNKESFWDFEPPEFTQSKY----LMIYKDNFTLDSNTSQIEVYLT 250

RESULT 45

AAW12146
 ID AAW12146 standard; Protein; 251 AA.
 XX AC AAW12146;
 XX 04-NOV-1997 (first entry)
 XX Streptococcus pyogenes Streptococcal toxin A mutant Lys157Glu.
 DE Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
 XX vaccine; protection; treatment; cancer; neutralising antibody;
 KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
 KW fever; hypotension; group A streptococcal infection; myositis;
 KW fasciitis; liver damage; T cell; lymphoma; ovarian; uterine.
 XX Streptococcus pyogenes.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Peptide 1..30
 FT /label= sig_peptide
 FT Peptide 31..251
 FT /label= mat_peptide
 FT Misc-difference 187
 FT /note= "wild type Lys replaced by Glu"
 XX WO9640930-A1.
 XX 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US10252.
 XX 07-JUN-1995; 95US-0480261.
 XX (MINU) UNIV MINNESOTA.
 XX Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
 XX WPI; 1997-099936/09.
 XX Mutant SPE-A toxin with at least one amino acid change is
 PT substantially non-lethal - used in vaccine composition for
 PT treatment of cancer and streptococcal toxic shock syndrome etc.
 XX Claim 5; Page -: 102pp; English.

XX The present sequence is a non-lethal Streptococcus pyogenes
 CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce
 CC vaccines to protect animals against wild type SPE-A and to treat
 CC cancer and streptococcal toxic shock syndrome (STSS). The mutant
 CC SPE-A causes neutralising antibodies (Ab) to be produced, which
 CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,
 CC group A streptococcal infection, myositis, fasciitis and liver
 CC damage. The neutralising Ab is preferably administered in
 CC conjunction with antibiotic therapy. The mutant SPE-A is
 CC especially useful for treating T cell lymphomas, and ovarian and
 CC uterine cancer. It is thought that mutant SPE-A can be selectively
 CC toxic to T cell lymphoma cells.
 CC N.B. Sequence not given in the specification, but constructed
 CC using the wild type SPE-A sequence given on pages 77-79.

XX Sequence 251 AA;

Query Match 22.0%; Score 268.5; DB 18; Length 251;
 Best Local Similarity 32.5%; Pred. No. 6.9e-19;
 Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;
 QY 4 SEINEKDLRKXSELOGTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLNTLLFGFFTG 60
 Db 25 SOEVAQQDPDSQLHRSSLVKNLQDIPLYEGDPVTHENVKSVQQLSHLLIYN--VS 81
 QY 61 HPWYNLLVDLGLSTAAATSEYEGSSVDLYGAYYGYOC-AGGTPNKTKACYGGVTLHDNNRL 119

Db 82 GPNYDKLTELKQEMATLFDKDNVDIYGVYHLYCENAEARSACIYGGVTNHEGNHL 141
 QY 120 TEEKVPIINLWIDGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFLGYNDSDFGGKV 179
 Db 142 EIPKKIVVKSVDIGIQ-SLSFD-IETNKKMVTQAQELDKVRKYLTDNEQLYTNGP--SKY 197
 QY 180 QRLIVFHSSEGSTVSVDLFD----AQOYPTDLLRIYRDNTTISSTLSISLYLYT 232
 Db 198 ETGYIKFIPKNKSEFWDFPEFTQSKY----LMIYKDNETLDSNTSQIEVYLT 250

RESULT 46
 AAW12097
 ID AAW12097 standard; Protein; 251 AA.

XX AC AAW12097;

XX DT 04-NOV-1997 (first entry)

XX DE Streptococcus pyogenes Streptococcal toxin A.

XX KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
 KW vaccine; protection; treatment; cancer; neutralising antibody;
 KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
 KW fever; hypotension; group A streptococcal infection; myositis;
 KW fasciitis; liver damage; T cell; lymphoma; ovarian; uterine.

XX OS Streptococcus pyogenes.

XX FH Key Location/Qualifiers
 FT Peptide 1..30

FT Peptide /label= sig_peptide

FT Peptide 31..251

XX FT /label= mat_peptide

XX PN W09640930-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US10252.

XX PR 07-JUN-1995; 95US-0480261.

XX PA (MINU) UNIV MINNESOTA.

XX PI Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;

XX WPI; 1997-099936/09.

XX DR N-PSDB; AAW12097.

XX PT Mutant SPE-A toxin with at least one amino acid change is

XX PT substantially non-lethal - used in vaccine composition for

XX PT treatment of cancer and streptococcal toxic shock syndrome etc.

XX PS Disclosure; Pages 77-79; 102pp; English.

XX CC The present sequence is Streptococcus pyogenes Streptococcal
 CC toxin A (SPE-A), from which a non-lethal mutant SPE-A, comprising
 CC at least 1 amino acid change, can be derived. The mutant SPE-A can
 CC be used to produce vaccines to protect animals against wild type
 CC SPE-A and to treat cancer and streptococcal toxic shock syndrome
 CC (STSS). The mutant SPE-A causes neutralising antibodies (Ab) to be
 CC produced, which may be used to ameliorate STSS symptoms, e.g.
 CC fever, hypotension, group A streptococcal infection, myositis,
 CC fasciitis and liver damage. The neutralising Ab is preferably
 CC administered in conjunction with antibiotic therapy. The mutant
 CC SPE-A is especially useful for treating T cell lymphomas, and
 CC ovarian and uterine cancer. It is thought that mutant SPE-A can be
 CC selectively toxic to T cell lymphoma cells.

XX SQ Sequence 251 AA;

Query Match 22.0%; Score 268.5; DB 18; Length 251;

Best Local Similarity 32.5%; Pred. No. 6.9e-19;
 Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SBEINEKDLRKSELOCTAL-CNLKQIYY-YNSKAIT-SSEKSADQFLTNTLLFKGFFTG 60
 Db 25 S0EVAQADPDPSQHRSSLVKNLQNIYFLYEGDPVTHENVKSVQDLSSHLLIYN--VS 81
 QY 61 HPWYNDLVLIDLGTAATSEYEGSSVDLYGAYGYOC-AGGTPNKTCACMYGGVTLHDNNRL 119
 Db 82 GPNYDKLTELKQEMATLFDKDNVDIYGVYHLYCENAEARSACIYGGVTNHEGNHL 141
 QY 120 TEEKVPIINLWIDGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFLGYNDSDFGGKV 179
 Db 142 EIPKKIVVKSVDIGIQ-SLSFD-IETNKKMVTQAQELDKVRKYLTDNKQLYTNGP--SKY 197
 QY 180 QRLIVFHSSEGSTVSVDLFD----AQOYPTDLLRIYRDNTTISSTLSISLYLYT 232
 Db 198 ETGYIKFIPKNKSEFWDFPEFTQSKY----LMIYKDNETLDSNTSQIEVYLT 250

RESULT 47

AAW59780

ID AAW59780 standard; Protein; 251 AA.

XX AC AAW59780;

XX DT 12-OCT-1998 (first entry)

XX DE Amino acid sequence of Streptococcus pyogenes exotoxin A.

XX KW SPE-A toxin; nonlethal; mutant; Streptococcus pyogenes exotoxin A;
 KW wild type; nontoxic; antibody; vaccine; immunity; ovarian cancer;
 KW streptococcal toxic shock syndrome; STSS; T cell lymphoma;
 KW uterine cancer.

XX OS Streptococcus pyogenes.

XX PN W09824911-A2.

XX PD 11-JUN-1998.

XX PF 05-DEC-1997; 97WO-US222228.

XX PR 06-DEC-1996; 96US-0032930.

XX PA (MINU) UNIV MINNESOTA.

XX PI Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;

XX WPI; 1998-333330/29.

XX DR N-PSDB; AAW41593.

XX PT New mutant Streptococcal SPE-A toxins - useful for, e.g. prevention
 or treatment of streptococcal infection or toxic shock syndrome

XX PS Disclosure; Fig 3; 95pp; English.

XX CC This is the amino acid sequence of the Streptococcus pyogenes exotoxin A
 CC (SPE-A toxin). The novel mutant Streptococcal SPE-A toxin has at least
 CC 1 aa change and is nonlethal compared with a protein to wild type SPE-A
 CC toxin. The mutant SPE-A toxins are nontoxic and can produce antibodies
 CC that neutralise wild type SPE-A toxin activity. The toxins can be used
 CC in vaccines and therapeutics to generate a protective immune response
 CC against streptococcal infection. They can be used to protect against the
 CC development of streptococcal toxic shock syndrome (STSS). In addition,
 CC the toxins can be used for treating animals with symptoms of
 CC streptococcal infection or STSS and in methods for stimulating T cell
 CC proliferation and in the treatment of cancer. In particular they can be
 CC used for treating T cell lymphomas, and ovarian and uterine cancer.

XX SQ Sequence 251 AA;

Query Match 22.0%; Score 268.5; DB 19; Length 251;


```
XX Sequence 239 AA;
SQ Query Match 22.0%; Score 267.5; DB 22; Length 239;
Best Local Similarity 31.7%; Pred. No. 8.1e-19;
Matches 77; Conservative 43; Mismatches 104; Indels 19; Gaps 8;

QY 2 EKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSEKSDAQPLTNTLLFKGFFTGH 61
D 1 ESQDPKPDDELHKSSKFTG-LMENMK-VLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKL 58
QY 62 PWNDLLDLGSGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTCMYGCV 111
D 59 GNYDNRVVEFKKDLADKYDKYDVDFGANYYYQCYFSKKTNDINSHQTDKRTCMYGV 118
QY 112 TLHDNNRLTEKKVPIINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYN 171
D 119 TEHNGNQLDKYRSITVRVFEDEGK-NLLSPD-VQTNKKVTAQELDYLTRHYLVKNNKLYE 176
QY 172 SDSFGKVGQGLIVFHSSEGSTVSVDLFAOQOYDP--TLLRIYRDNNTTISSTLSISLY 229
D 177 FNN--SPYETGYIKFIENENS-FWYDMPAPGDKFQDSKYLMMYNDKNKWDKDKYKIEVY 233
QY 230 LYT 232
D 234 LTT 236

RESULT 50
AAW06737
ID AAW06737 standard; Protein; 255 AA.
XX
AC AAW06737;
DT 08-MAR-1997 (first entry)
XX
DE Staphylococcus enterotoxin B.
XX
KW Enterotoxin B; superantigen; antigen; cytokine; chemokine; T cell;
KW lymphocyte; monocyte; natural killer cell; gene therapy; cancer;
KW vaccine; adjuvant.
XX
OS Staphylococcus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..15
FT /label= Sig_peptide
XX
PN W09G36366-A1.
XX
PD 21-NOV-1996.
XX
PF 20-MAY-1996; 96WO-US07432.
XX
PR 29-DEC-1995; 95US-0580806.
PR 18-MAY-1995; 95US-0446918.
XX
PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY MED.
XX
PI Dow SW, Elmslie RE, Potter TA;
XX
WPI; 1997-011857/01.
DR N-PSDB; AAT45698.
XX
PT Recombinant molecule encoding superantigen and opt. cytokine or
PT chemokine - controls activity of effector cells (T cells, monocytes,
PT natural killer cells), used for gene therapy of cancer
XX
PS Example 1; Page 96-97; 131pp; English.
XX
CC A cDNA clone (AAT45698) codes for staphylococcal enterotoxin B
CC (AAW06737) superantigen. Nucleic acids encoding superantigens (see
CC also AAW06738-39), esp. truncated forms of the superantigen lacking
```

```
CC the leader peptide, can be utilised in the gene therapy of cancer,
CC infectious diseases and immunological disorders. The nucleic acid,
CC optionally in combination with cytokine or chemokine nucleic acids,
CC is delivered to an animal using e.g. liposomes. It acts by
CC controlling the activity of effector cells, such as T-cells,
CC macrophages, monocytes and/or natural killer cells. Localised
CC prodn. of an effective but non-toxic amount of encoded proteins
CC allows safe treatment of the animal.
XX
SQ Sequence 255 AA;

Query Match 22.0%; Score 267.5; DB 18; Length 255;
Best Local Similarity 31.7%; Pred. No. 8.9e-19;
Matches 77; Conservative 43; Mismatches 104; Indels 19; Gaps 8;

QY 2 EKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSEKSDAQPLTNTLLFKGFFTGH 61
D 17 ESQDPKPDDELHKSSKFTG-LMENMK-VLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKL 74
QY 62 PWNDLLDLGSGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTCMYGCV 111
D 75 GNYDNRVVEFKKDLADKYDKYDVDFGANYYYQCYFSKKTNDINSHQTDKRTCMYGV 134
QY 112 TLHDNNRLTEKKVPIINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYN 171
D 135 TEHNGNQLDKYRSITVRVFEDEGK-NLLSPD-VQTNKKVTAQELDYLTRHYLVKNNKLYE 192
QY 172 SDSFGKVGQGLIVFHSSEGSTVSVDLFAOQOYDP--TLLRIYRDNNTTISSTLSISLY 229
D 193 FNN--SPYETGYIKFIENENS-FWYDMPAPGDKFQDSKYLMMYNDKNKWDKDKYKIEVY 249
QY 230 LYT 232
D 250 LTT 252

Search completed: January 8, 2003, 11:56:54
Job time : 21.4923 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: January 8, 2003, 11:55:37 ; Search time 6.95138 Seconds
(without alignments)
986.213 Million cell updates/sec

Title: US-09-900-766-2

Perfect score: 1218

Sequence: 1 SEKSEEINEXDLRKKSELOG.....RDNTTISISLSISLYLTT 233

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1107	90.9	257	3	US-08-486-099-112
2	1107	90.9	257	3	US-08-360-107A-112
3	1107	90.9	257	3	US-08-484-223B-112
4	1107	90.9	257	3	US-08-919-597-112
5	1107	90.9	257	3	US-08-475-668A-112
6	1107	90.9	257	3	US-08-485-551A-112
7	1107	90.9	257	3	US-08-471-913A-112
8	1107	90.9	257	4	US-08-485-264A-112
9	1107	90.9	257	4	US-08-474-349A-112
10	1040	85.4	226	4	US-08-896-933-24
11	1040	85.4	226	4	US-09-314-235-24
12	948	77.8	257	3	US-08-486-099-113
13	948	77.8	257	3	US-08-360-107A-123
14	948	77.8	257	3	US-08-484-223B-113
15	948	77.8	257	3	US-08-919-597-113
16	948	77.8	257	3	US-08-475-668A-113
17	948	77.8	257	3	US-08-485-551A-113
18	948	77.8	257	3	US-08-471-913A-113
19	948	77.8	257	4	US-08-485-264A-113
20	948	77.8	257	4	US-08-474-349A-113
21	944	77.5	233	1	US-08-446-918A-4
22	944	77.5	233	2	US-08-580-806-4
23	931.5	76.5	232	4	US-08-896-933-23
24	931.5	76.5	232	4	US-09-314-235-23
25	925	75.9	257	4	US-09-144-776B-2
26	921	75.6	233	4	US-09-144-776B-4
27	605	49.7	228	4	US-08-896-933-25

28	605	49.7	228	4	US-09-314-235-25
29	406	33.3	82	4	US-09-144-776B-19
30	362	29.7	82	4	US-09-144-776B-17
31	277.5	22.8	238	4	US-08-896-933-28
32	277.5	22.8	238	4	US-09-314-235-28
33	270.5	22.2	221	4	US-08-896-933-29
34	270.5	22.2	221	4	US-09-314-235-29
35	267.5	22.0	255	1	US-08-446-918A-2
36	267.5	22.0	255	2	US-08-580-806-2
37	264.5	21.7	239	4	US-08-896-933-27
38	264.5	21.7	239	4	US-09-314-235-27
39	263.5	21.6	251	4	US-09-144-776B-16
40	260.5	21.4	239	4	US-08-896-933-26
41	260.5	21.4	239	4	US-09-314-235-26
42	260.5	21.4	266	4	US-09-414-276-8
43	260.5	21.4	266	4	US-09-144-776B-6
44	255.5	21.0	239	4	US-09-144-776B-10
45	255.5	21.0	266	4	US-09-144-776B-14

ALIGNMENTS

RESULT 1

US-08-486-099-112

; Sequence 112, Application US/08486099

; Patent No. 6013263

; GENERAL INFORMATION:

; APPLICANT: Bolognesi, Dani P.

; APPLICANT: Matthews, Thomas J.

; APPLICANT: Wild, Carl T.

; APPLICANT: Barney, Shawn O.

; APPLICANT: Lambert, Dennis M.

; APPLICANT: Petteway, Stephen R.

; APPLICANT: Langlois, Alphonse J.

; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF

; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS

; NUMBER OF SEQUENCES: 209

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,099

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7872-031

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEFAX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 112:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 257 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; US-08-486-099-112

Query Match 90.9%; Score 1107; DB 3; Length 257;

Best Local Similarity	89.78;	prec	No. 1.6e-110;						
Matches	209;	Conservative	9;	Mismatches	15;	Indels		0;	Gaps
Qy	1	SEKSEINEKDLRKKSELQGTALGNLKIYYIYNKSAITSEKSDAQFTNTLLFKGFFTG	60						
Db	25	SEKSEINEKDLRKKSELQRNALSRLRIYYIYNKSAITENKSDSDQFLENTLLFKGFFTG	84						
Qy	61	HPWYNLLLDLGSAAATSEYEGSSVDLYGAYVGYOCAGTGNKTKACMYGGVTLHDNNRLT	120						
Db	85	HPWYNLLLDLGSKDATNKYGGKVDLYGAYVGYOCAGTGNKTKACMYGGVTLHDNNRLT	144						
Qy	121	EKKVPINLWIDGQTTVPIDKVTSTKEVTVQELDLQARHLYHGKFGLYNSDSFGGKVQ	180						
Db	145	EKKVPINLWIDGQTTVPIDKVTSTKEVTVQELDLQARHLYHGKFGLYNSDSFGGKVQ	204						
Qy	181	RGLVFIHSSSEGSTVSYDLFDAGQGYPDPLLRIRYRDNTTISSTSSISISLYLXTT	233						
Db	205	RGLVFIHSSSEGSTVSYDLFDAGQGYPDPLLRIRYRDNTKINSNLHLDLYLYTT	257						

```

RESULT 2
US-08-360-107A-122
; Sequence 122, Application US/08360107A
; Patent No. 6017536
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,107A
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-360-107A-122

```

```

RESULT 3
US-08-484-223B-112
> Sequence 112, Application US/08484223B
> Patent No. 6020459
> GENERAL INFORMATION:
> APPLICANT: Bolognesi, Dani P.
> APPLICANT: Matthews, Thomas J.
> APPLICANT: Wild, Carl T.
> APPLICANT: Barney, Shawn O.
> APPLICANT: Lambert, Dennis M.
> APPLICANT: Petteway, Stephen R.
> APPLICANT: Langlois, Alphonse J.
> TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
> TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
> TITLE OF INVENTION: TRANSMISSION
> NUMBER OF SEQUENCES: 245
> CORRESPONDENCE ADDRESS:
> ADDRESSEE: Pennie & Edmonds LLP
> STREET: 1155 Avenue of the Americas
> CITY: New York
> STATE: New York
> COUNTRY: USA
> ZIP: 10036-2711
> COMPUTER READABLE FORM:
> MEDIUM TYPE: Floppy disk
> COMPUTER: IBM PC compatible
> OPERATING SYSTEM: PC-DOS/MS-DOS
> SOFTWARE: PatentIn Release #1.0, Version #1.30
> CURRENT APPLICATION DATA:
> APPLICATION NUMBER: US/08/484,223B
> FILING DATE: 07-JUN-1995
> CLASSIFICATION: 435
> ATTORNEY/AGENT INFORMATION:
> NAME: Coruzzi, Laura A.
> REGISTRATION NUMBER: 30,742
> REFERENCE/DOCKET NUMBER: 7872-029
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: (212) 790-9090
> TELEFAX: (212) 869-9741/8864
> TELEX: 66141 PENNIE
> INFORMATION FOR SEQ ID NO: 112:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 257 amino acids
> TYPE: amino acid
> STRANDEDNESS:
> TOPOLOGY: unknown
> MOLECULE TYPE: protein
US-08-484-223B-112

```

```

Query Match      90.9%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 SEKSEENEKDLRKSELGQTALGNLKIQIYYNKSKAITSSEKSAQDFLNTLLPKGFETG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 SEKSEENEKDLRKSELRNALSNLROIYYNKSKAITENKESDDQFLENTLLPKGFETG 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Qy	61	HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKATCMYGGVTLHDNNRLT	120
Db	85	HPWYNDLLVDLGSKDATNKYKGKKVDLYGAYGYQCAGGTPNKATCMYGGVTLHDNNRLT	144
Qy	121	EKKVPINLWDGKQTTPIDKVKTSKKEVTVQELDQARHYLHGKFGLYNSDSFGGKYQ	180
Db	145	EKKVPINLWDGKQTTPIDKVKTSKKEVTVQELDQARHYLHGKFGLYNSDSFGGKYQ	204
Qy	181	RGLIVFHSSGGSVSDYDPAQOQYPTLLRIYRDNTTISSTLSLSLYLT	233
Db	205	RGLIVFHSSGGSVSDYDPAQOQYPTLLRIYRDNTTINSENLDLYLT	257

RESULT 4
US-08-919-597-112
: Sequence 112, Application US/08919597
: Patent No. 6054265
: GENERAL INFORMATION:
: APPLICANT: Bolognesi, Dani P.
: APPLICANT: Matthews, Thomas J.
: APPLICANT: Wild, Carl T.
: APPLICANT: Barney, Shawn O.
: APPLICANT: Lambert, Dennis M.
: APPLICANT: Petteway, Stephen R.
: APPLICANT: Langlois, Alphonse J.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
: TITLE OF INVENTION: TRANSMISSION
: NUMBER OF SEQUENCES: 273
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennle & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/919,597
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/470,896
: FILING DATE: 06-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7872-020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 112:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 257 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
US-08-919-597-112

```

Qy 61 HPWYNDLLVLDGTAATSEYEGSSVDLYGAYGYQACAGTGNKTKTACMYGGVTLHDNNRLT 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 85 HPWYNDLLVLDGSKDATNKYKGGKVDLYGAYGYQACAGTGNKTKTACMYGGVTLHDNNRLT 144
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 121 EEKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHCKFGLYNSDSFGGKVQ 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 145 EEKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHCKFGLYNSDSFGGKVQ 204
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy 181 RGLVHSSSEGSTSVSYDLFDAQQGYPDFTLLRIYRDNTTISSTLSISLYLYTT 233
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 205 RGLVHSSSEGSTSVSYDLFDAQQGYPDFTLLRIYRDNTTINSENHLIDLILYLYTT 257
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 5
US-08-475-668A-112
; Sequence 112, Application US/08475668A
; Patent No. 6060065
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petceway, Stephen R.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds LLP.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475.668A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-475-668A-112

```

	Query Match	90.9%	Score 1107;	DB 3;	Length 257;
	Best Local Similarity	89.7%;	Pred. No. 1.6e-110;		
	Matches 209;	Conservative 9;	Mismatches 15;	Indels 0;	Gaps 0;
Qy	1	SEKSEEINEKDLRKKSELQGTALGNLKOIYYINYNSKAITSSSEKSADQFUTNTLLPKGFPTG	60		
Db	25	SEKSEEINEKDLRKKSELQRNALSNLROIYYINKEAITENKESDQFLENTLLPKGFPTG	84		
Qy	61	HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMVGVTLDHNNRLT	120		
		: :			
Db	85	HPWYNDLLVDLGSKDATNKKYGKKVDLYGAYGYQCAGGTPNKTCMVGVTLDHNNRLT	144		
Qy	121	EKKVPINLWDIGKQTTPIDIKVTKSKEVTQVELDLQARHYLHGKFGLYNSDSFGGVKVQ	180		

```
Db 145 EEKVPINLWIDGKQTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGTSVSYDLFDAOGQYPTDLLRIYRDNNTTISSTSLSLYLYTT 233
Db 205 RGLIVFHSSEGTSVSYDLFDAOGQYPTDLLRIYRDNKTINSENHLIDLTYTT 257

RESULT 6
US-08-485-551A-112
; Sequence 112, Application US/08485551A
; Patent No. 6068973
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,551A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-551A-112

Query Match 90.9%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELQGTALGNLKQIYYNYSKAITSEKSAQOFLNTLLFKGFFTG 60
Db 25 SEKSEINEKDLRKSELQARNLSNLQIYYNYSKAITENKESDDQFLNTLLFKGFFTG 84
QY 61 HPWYNDLLVDLGSTATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGCVTLHDNNRLT 120
Db 85 HPWYNDLLVDLGSKATNRYKGGKVDLYGAYGYQCAGGTPNKTACMYGCVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
Db 145 EEKVPINLWIDGKQTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGTSVSYDLFDAOGQYPTDLLRIYRDNNTTISSTSLSLYLYTT 233
QY 181 RGLIVFHSSEGTSVSYDLFDAOGQYPTDLLRIYRDNKTINSENHLIDLTYTT 257
```

```
Db 205 RGLIVFHSSEGTSVSYDLFDAOGQYPTDLLRIYRDNKTINSENHLIDLTYTT 257

RESULT 7
US-08-471-913A-112
; Sequence 112, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,913A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-471-913A-112

Query Match 90.9%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELQGTALGNLKQIYYNYSKAITSEKSAQOFLNTLLFKGFFTG 60
Db 25 SEKSEINEKDLRKSELQARNLSNLQIYYNYSKAITENKESDDQFLNTLLFKGFFTG 84
QY 61 HPWYNDLLVDLGSTATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGCVTLHDNNRLT 120
Db 85 HPWYNDLLVDLGSKATNRYKGGKVDLYGAYGYQCAGGTPNKTACMYGCVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
Db 145 EEKVPINLWIDGKQTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGTSVSYDLFDAOGQYPTDLLRIYRDNNTTISSTSLSLYLYTT 233
Db 205 RGLIVFHSSEGTSVSYDLFDAOGQYPTDLLRIYRDNKTINSENHLIDLTYTT 257
```

RESULT 8

US-08-485-264A-112
; Sequence 112, Application US/08485264A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.264A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-264A-112

Query Match 90.9%; Score 1107; DB 4; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;
QY 1 SEKSEINEKDLRKKSELOGTALGNLKOIYYNNKAITSSSEKSADQFLTNLLFKGFTG 60
DB 25 SEKSEINEKDLRKKSELQNALSNLROIYYNNKAITENKESDDQFLTNLLFKGFTG 84
QY 61 HPWYNDLLVLDLSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
DB 85 HPWYNDLLVLDLSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
DB 145 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTT 233
DB 205 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTT 257

RESULT 9

US-08-474-349A-112

; Sequence 112, Application US/08474349A
; Patent No. 6333395
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
; TITLE OF INVENTION: VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 517
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474.349A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-474-349A-112

Query Match 90.9%; Score 1107; DB 4; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;
QY 1 SEKSEINEKDLRKKSELOGTALGNLKOIYYNNKAITSSSEKSADQFLTNLLFKGFTG 60
DB 25 SEKSEINEKDLRKKSELQNALSNLROIYYNNKAITENKESDDQFLTNLLFKGFTG 84
QY 61 HPWYNDLLVLDLSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
DB 85 HPWYNDLLVLDLSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
DB 145 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTT 233
DB 205 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTT 257

RESULT 10

US-08-896-933-24
; Sequence 24, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:

APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
SUPERANTIGENS, AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005002
CURRENT APPLICATION NUMBER: US/08/896.933
EARLIER FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252.978
EARLIER FILING DATE: 1994-06-02
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 226
TYPE: PRT
ORGANISM: Staphylococcus aureas
US-08-896-933-24

Query Match 85.4%; Score 1040; DB 4; Length 226;
Best Local Similarity 85.2%; Pred. No. 2e-103;
Matches 196; Conservative 13; Mismatches 17; Indels 4; Gaps 1;
QY 4 SEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTGHPW 63
DB 1 SEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTGHPW 60
QY 64 YNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEK 123
DB 61 YNDLLVDKGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEE- 119
QY 124 KVPINLWDGKOTTPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGL 183
DB 120 ---VBKWIDGKOTTPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGL 176
QY 184 IVFHSSEGSTVSYDLFDAQQGYPTDLLRIYRDNNTTISSTLSISLYLYTT 233
DB 177 IVFHSSEGSTVSYDLFDAQQGYPTDLLRIYRDNNTTINSENHIDIYLYTT 226

RESULT 11
US-09-314-235-24
Sequence 24, Application US/09314235
Patent No. 6338845
GENERAL INFORMATION:
APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
SUPERANTIGENS, AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005004
CURRENT APPLICATION NUMBER: US/09/314,235
CURRENT FILING DATE: 1999-05-18
EARLIER APPLICATION NUMBER: 08/896.933
EARLIER FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252.978
EARLIER FILING DATE: 1994-06-02
EARLIER APPLICATION NUMBER: 07/891.718
EARLIER FILING DATE: 1992-06-01
EARLIER APPLICATION NUMBER: US91/00342
EARLIER FILING DATE: 1991-01-17
EARLIER APPLICATION NUMBER: 07/466.577
EARLIER FILING DATE: 1990-01-17
EARLIER APPLICATION NUMBER: 07/416.530
EARLIER FILING DATE: 1989-10-03
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 226
TYPE: PRT
ORGANISM: Staphylococcus aureas
US-09-314-235-24

Query Match 85.4%; Score 1040; DB 4; Length 226;
Best Local Similarity 85.2%; Pred. No. 2e-103;
Matches 196; Conservative 13; Mismatches 17; Indels 4; Gaps 1;
QY 4 SEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTGHPW 63

DB 1 SEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTGHPW 60
QY 64 YNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEK 123
DB 61 YNDLLVDKGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEE- 119
QY 124 KVPINLWDGKOTTPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGL 183
DB 120 ---VBKWIDGKOTTPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGL 176
QY 184 IVFHSSEGSTVSYDLFDAQQGYPTDLLRIYRDNNTTISSTLSISLYLYTT 233
DB 177 IVFHSSEGSTVSYDLFDAQQGYPTDLLRIYRDNNTTINSENHIDIYLYTT 226

RESULT 12
US-08-486-099-113
Sequence 113, Application US/08486099
Patent No. 6013263
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
TITLE OF INVENTION: B VIRUS TRANSMISSION
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,099
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-031
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-486-099-113

Query Match 77.8%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 1.7e-93;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;
QY 1 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTG 60
DB 25 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTG 84


```
RESULT 17
US-08-485-551A-113
; Sequence 113, Application US/08485551A
; Patent No. 6068973
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,551A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 869-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-485-551A-113

Query Match 77.8%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 1.7e-93;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKKSELOGTALGNLKOIYYNKAITSSEKSADOFLLTLLFKGFFTG 60
Db 25 SEKSEINEKDLRKKSELOGTALGNLKOIYYNKAITSSEKSADOFLLTLLFKGFFTD 84
Qy 61 HPWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 85 HSWYNLLVDFDSKDIDYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
Qy 121 EEKKVPINLWIDGKQTPVIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQ 180
Db 145 EEKKVPINLWLDGKQNTVPLETVKTNKKNVTVOELDLQARRYLQEKYLNLYNSDVFQKQV 204
Qy 181 RGLIVPHSSEGSTSVSYDLFDAQOGYPTDLLRIYRONTTISSTLSLSLYTT 233
Db 205 RGLIVPHTSTEPSVNYDLFGAOGQVSNLTLLRIYRONKNTINSENMHIDIYLYTS 257

RESULT 18
US-08-471-913A-113
; Sequence 113, Application US/08471913A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
```

```
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,264A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-021
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: protein
; MOLECULE TYPE: protein
; US-08-485-264A-113

Query Match 77.8%; Score 948; DB 4; Length 257;
Best Local Similarity 76.4%; Pred. No. 1.7e-93;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTLTLFKGFFTG 60
Db 25 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTLTLFKGFFTG 84
Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOCAGTPNKTACMYGGVTLHDNNRLT 120
Db 85 HSWYNDLLVDFDSKDIVDKYKGVVLYGAYGYOCAGTPNKTACMYGGVTLHDNNRLT 144
Qy 121 EEKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLGYNLSDSFGGKVQ 180
Db 145 EEKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLGYNLSDSFGGKVQ 204
Qy 181 RGLIVFHSSEGSTSVYDLFDAGQGYPTDRLRIYRDNKTINSNMHIDIYLYTS 233
Db 205 RGLIVFHSSEGSTSVYDLFDAGQGYPTDRLRIYRDNKTINSNMHIDIYLYTS 257

RESULT 20
US-08-474-349A-113
; Sequence 113, Application US/08474349A
; Patent No. 6333395
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Danl P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
```

```
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
; TITLE OF INVENTION: VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 517
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,349A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-024
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: protein
; MOLECULE TYPE: protein
; US-08-474-349A-113

Query Match 77.8%; Score 948; DB 4; Length 257;
Best Local Similarity 76.4%; Pred. No. 1.7e-93;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTLTLFKGFFTG 60
Db 25 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTLTLFKGFFTG 84
Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOCAGTPNKTACMYGGVTLHDNNRLT 120
Db 85 HSWYNDLLVDFDSKDIVDKYKGVVLYGAYGYOCAGTPNKTACMYGGVTLHDNNRLT 144
Qy 121 EEKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLGYNLSDSFGGKVQ 180
Db 145 EEKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLGYNLSDSFGGKVQ 204
Qy 181 RGLIVFHSSEGSTSVYDLFDAGQGYPTDRLRIYRDNKTINSNMHIDIYLYTS 233
Db 205 RGLIVFHSSEGSTSVYDLFDAGQGYPTDRLRIYRDNKTINSNMHIDIYLYTS 257

RESULT 21
US-08-446-918A-4
; Sequence 4, Application US/08446918A
; Patent No. 5705151
; GENERAL INFORMATION:
; APPLICANT: Dow, Steve W.
; APPLICANT: Elmslie, Robyn E.
; TITLE OF INVENTION: GENE THERAPY FOR T CELL REGULATION
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
```

CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,918A
FILING DATE: 18-MAY-1995
CLASSIFICATION: 552
ATTORNEY/AGENT INFORMATION:
NAME: Kovarik, Joseph E.
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2879-29
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-918A-4

Query Match 77.5%; Score 944; DB 1; Length 233;
Best Local Similarity 76.3%; Pred. No. 4e-93;
Matches 177; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 2 EKSEINEKDLRKKSELOGTALGNLKOIYYYNKAITTSSEKSADQFLTNTLLFKGFFTG 61
Db 2 EKSEINEKDLRKKSELOGTALGNLKOIYYYNKAKTENKESHQFLQHTILFKGFFTDH 61

QY 62 PWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
Db 62 SWYNLLVDFDSKDIDVKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121

QY 122 EKKVPLNLWDGKQTPVIDKYKSKKVTVOELDLQARHYLHGKFGLYNSDSFGKQVOR 181
Db 122 EKKVPLNLWDGKQTPVLETVKTKNKNVTVOELDLQARRYLQEKYLYNSDVFQKQVOR 181

QY 182 GLIVFHSSSGSVSYDLFDAQOQYPTLLRIYRDNTTISSTLSISLYLYTT 233
Db 182 GLIVFHTSTPSVNYDLFGAQOYSNTLLRIYRDNTINSENHIDIYLYTS 233

RESULT 22
US-08-580-806-4
Sequence 4, Application US/08580806
Patent No. 5935568
GENERAL INFORMATION:
APPLICANT: Dow, Steve W.
APPLICANT: Elmslie, Robyn E.
APPLICANT: Potter, Terence A.
TITLE OF INVENTION: GENE THERAPY FOR EFFECTOR CELL REGULATION
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/580,806

FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-29-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-580-806-4

Query Match 77.5%; Score 944; DB 2; Length 233;
Best Local Similarity 76.3%; Pred. No. 4e-93;
Matches 177; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 2 EKSEINEKDLRKKSELOGTALGNLKOIYYYNKAITTSSEKSADQFLTNTLLFKGFFTG 61
Db 2 EKSEINEKDLRKKSELOGTALGNLKOIYYYNKAKTENKESHQFLQHTILFKGFFTDH 61

QY 62 PWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
Db 62 SWYNLLVDFDSKDIDVKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121

QY 122 EKKVPLNLWDGKQTPVIDKYKSKKVTVOELDLQARHYLHGKFGLYNSDSFGKQVOR 181
Db 122 EKKVPLNLWDGKQTPVLETVKTKNKNVTVOELDLQARRYLQEKYLYNSDVFQKQVOR 181

QY 182 GLIVFHSSSGSVSYDLFDAQOQYPTLLRIYRDNTTISSTLSISLYLYTT 233
Db 182 GLIVFHTSTPSVNYDLFGAQOYSNTLLRIYRDNTINSENHIDIYLYTS 233

RESULT 23
US-08-896-933-23
Sequence 23, Application US/08896933
Patent No. 6221351
GENERAL INFORMATION:
APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005002
CURRENT APPLICATION NUMBER: US/08/896,933
CURRENT FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252,978
EARLIER FILING DATE: 1994-06-02
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 232
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-896-933-23

Query Match 76.5%; Score 931.5; DB 4; Length 232;
Best Local Similarity 76.0%; Pred. No. 8.7e-92;
Matches 177; Conservative 21; Mismatches 34; Indels 1; Gaps 1;

QY 1 SEKSEINEKDLRKKSELOGTALGNLKOIYYYNKAITTSSEKSADQFLTNTLLFKGFFTG 60
Db 1 SEKSEINEKDLRKKSELOGTALGNLKOIYYYNKAKTENKESHQFL-HTILFKGFFTD 59

QY 61 HPWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 60 HSWYNLLVDFDSKDIDVKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 119

QY 121 EKKVPLNLWDGKQTPVIDKYKSKKVTVOELDLQARHYLHGKFGLYNSDSFGKQV 180
Db 121 EKKVPLNLWDGKQTPVLETVKTKNKNVTVOELDLQARRYLQEKYLYNSDVFQKQV 180

```
Db 120 EBKVPINLWLDGKQNTVPLETVTKNKNVTQVELDPOARRYLQEKYNLYNSDVFQKQV 179
QY 181 RGLIVFHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNNTTISSTLSISLYLYTT 233
|||||1 -||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 RGLIVFHTSTEPSVNDLFGAQOQYNTLLRIYRDNKNTINSENMMHIDIYLYTS 232

RESULT 24
US-09-314-235-23
; Sequence 23, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; CURRENT FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416,530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-23

Query Match 76.5%; Score 931.5; DB 4; Length 232;
Best Local Similarity 76.0%; Pred. No. 8.7e-92;
Matches 177; Conservative 21; Mismatches 34; Indels 1; Gaps 1;

QY 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTG 60
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTG 59
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 61 HPWYNDLVLDGTAATSEYEGSSVDLYGAYGYOCAGTGNKTCMYGGVTLHDNNRLT 120
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 60 HSWYNDLVDFDSKDIVDKYKGYDLYGAYGYOCAGTGNKTCMYGGVTLHDNNRLT 119
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 121 EEKVPINLWIDGKQNTVPIDKVTSKKEVTVOELDQARHLYHGKFGLYNSDSFGKQV 180
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 120 EEKVPINLWIDGKQNTVPLETVTKNKNVTQVELDPOARRYLQEKYNLYNSDVFQKQV 179
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 181 RGLIVFHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNNTTISSTLSISLYLYTT 233
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 180 RGLIVFHTSTEPSVNDLFGAQOQYNTLLRIYRDNKNTINSENMMHIDIYLYTS 232
```

```
RESULT 25
US-09-144-776B-2
; Sequence 2, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
; Mark A. Olson
; Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
; Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army MRC - 504 Scott Street
```

```
; MCMR-JA (Charles H. Harris-Patent
; Atty)
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,776B
; FILING DATE: 01-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/882,431
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-144-776B-2

Query Match 75.9%; Score 925; DB 4; Length 257;
Best Local Similarity 75.1%; Pred. No. 5e-91;
Matches 175; Conservative 21; Mismatches 37; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTG 50
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 25 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTG 84
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 61 HPWYNDLVLDGTAATSEYEGSSVDLYGAYGYOCAGTGNKTCMYGGVTLHDNNRLT 120
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 85 HSWYNDLVDFDSKDIVDKYKGYDLYGAYGYOCAGTGNKTCMYGGVTLHDNNRLT 144
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 121 EEKVPINLWIDGKQNTVPIDKVTSKKEVTVOELDQARHLYHGKFGLYNSDSFGKQV 180
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 145 EEKVPINLWIDGKQNTVPLETVTKNKNVTQVELDQARRYLQEKYNLYNSDVFQKQV 204
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 181 RGLIVFHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNNTTISSTLSISLYLYTT 233
|||||1 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 205 RGLIVFHTSTEPSVNDLFGAQOQYNTLLRIYRDNKNTINSENMMHIDIYLYTS 257

RESULT 26
US-09-144-776B-4
; Sequence 4, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
; Mark A. Olson
; Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
; Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army MRC - 504 Scott Street
; MCMR-JA (Charles H. Harris-Patent
; Atty)
; CITY: FORT DETRICK
```

```

; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,776B
; FILING DATE: 01-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/882,431
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-144-776B-4

Query Match 75.6%; Score 921; DB 4; Length 233;
Best Local Similarity 75.0%; Pred. No. 1.2e-90;
Matches 174; Conservative 21; Mismatches 37; Indels 0; Gaps 0;

QY 2 EKSEENKDLRKSELOGTALGNLQIYYNSKAITSEKSADQFLTNLLFKGFFTG 61
Db 2 EKSEENKDLRKSELOGTALGNLQIYYNEKAKTENKESHDPQHTILFKGFFTDH 61

QY 62 PWNDLLVLDGTAATSEYEGSSVDLYGAYGYOCAGGTPNKTCACMYGGVTLHDNNRLTE 121
Db 62 SWNDLLVRFDSKDIYDKYKGVLDLYGAYGOCAGGTPNKTCACMYGGVTLHDNNRLTE 121

QY 122 EKKVPINLWDGKQTTVPIDKVTYSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQR 181
Db 122 EKKVPINLWDGKQNTVPLETVTKTKNNTVOELDLQARRYLQEKYLYNSDVFDCGVQR 181

QY 182 GLIVFHSSEGSTVSYDLFDAOGQYPTLLRIYRDNTTISLSISLYLTT 233
Db 182 GLIVFHTSTEPSVNYDLFQAQGYSTLLRIYRDNTINSENHIDIYLYTS 233

RESULT 27
US-08-896-933-25
; Sequence 25, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-25

Query Match 49.7%; Score 605; DB 4; Length 228;
Best Local Similarity 52.0%; Pred. No. 7.9e-57;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;

QY 7 INEKDLRKSELOGTALGNLQIYYNSKAITSEKSADQFLTNLLFKGFFTGHPWYND 66
Db 2 VKEKELHKSELSSTALNNMKHSYADKNPIGENKSTGDOFLENTLLYKFFTDLINFED 61

QY 67 LLVDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTCACMYGGVTLHDNNRLTEKKVP 126
Db 62 LLINFNSKEMAQHFHFKSNVDVPIRISINCYGGEIDRTACTYGGVTPHEGNKLKERKKIP 121

QY 127 INLWIDGKQTTVPIDKVTYSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGLIVF 186
Db 122 INLWINGVQKEVSLDKVQTDKKNVTVOELDLQARRYLQKDLKLYNNDTLGGKTQKRGKIEF 181

QY 187 HSSEGSTVSYDLFDAOGQYPTLLRIYRDNTTISLSISLYLY 231
Db 182 DSSDGSKVSYDLFDVKGDFPEKQRIYSDNKTLSLSTELHIDIYLY 226

RESULT 28
US-09-314-235-25
; Sequence 25, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; EARLIER FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416,530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-25

Query Match 49.7%; Score 605; DB 4; Length 228;
Best Local Similarity 52.0%; Pred. No. 7.9e-57;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;

QY 7 INEKDLRKSELOGTALGNLQIYYNSKAITSEKSADQFLTNLLFKGFFTGHPWYND 66
Db 2 VKEKELHKSELSSTALNNMKHSYADKNPIGENKSTGDOFLENTLLYKFFTDLINFED 61

QY 67 LLVDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTCACMYGGVTLHDNNRLTEKKVP 126
Db 62 LLINFNSKEMAQHFHFKSNVDVPIRISINCYGGEIDRTACTYGGVTPHEGNKLKERKKIP 121

QY 127 INLWIDGKQTTVPIDKVTYSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGLIVF 186
Db 122 INLWINGVQKEVSLDKVQTDKKNVTVOELDLQARRYLQKDLKLYNNDTLGGKTQKRGKIEF 181

QY 187 HSSEGSTVSYDLFDAOGQYPTLLRIYRDNTTISLSISLYLY 231
Db 182 DSSDGSKVSYDLFDVKGDFPEKQRIYSDNKTLSLSTELHIDIYLY 226

US-08-896-933-25
```

[illegible]


```
Query Match      22.2%; Score 270.5; DB 4; Length 221;
Best Local Similarity 33.8%; Pred. No. 5.1e-21;
Matches 76; Conservative 42; Mismatches 28; Indels 19; Gaps 10;

Qy 16 SELOQTAL-GNLKQIYV-VNSKAIT--SSPKSADQFLTNLLFKGFFTHGPWYNLLVDLG 72
Db 7 SOLHRSSLVKNLQNIYFLYEGDPVTHENVKSVQDQLLSHDLIYN---VSGPNYDKLKTDL 63

Qy 73 STAATSEYEGSSVDLYGAYGYQC-AGGTPNKATACMYGGVTLHDNNRLTEKKVPINLWI 131
Db 64 NQEMATLPKDKNVDIYGVYEHLCYLCEAERSACIYGGVTNHEGNHLEIPKKIVKYSI 123

Qy 132 DSKQTPVDIKVKTSSKEVTVQELDLQARHYLHGKGLYNSDFGKVGKVGRLIVFHSSEG 191
Db 124 DGIO-SLSFD-IETNKKMVTAEQELDYKVRKYLTDNKNQLYTNGP--SKYETGYIKFIPK 179

Qy 192 STVSVDLDF---ACQGYPTDLRLIYRDNTTISSTLSLSLYLT 232
Db 180 ESEWDLFPPEPTQSKY----LMYKDNELDSNTSQIEVILTT 220

RESULT 35
US-08-446-918A-2
; Sequence 2, Application US/08446918A
; Patent No. 5705151
; GENERAL INFORMATION:
; APPLICANT: Dow, Steve W.
; APPLICANT: Elmslie, Robyn E.
; TITLE OF INVENTION: GENE THERAPY FOR T CELL REGULATION
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.918A
; FILING DATE: 18-MAY-1995
; CLASSIFICATION: 552
; ATTORNEY/AGENT INFORMATION:
; NAME: Kovarik, Joseph E.
; REGISTRATION NUMBER: 33,005
; REFERENCE/DOCKET NUMBER: 2879-29
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-446-918A-2

Query Match      22.0%; Score 267.5; DB 1; Length 255;
Best Local Similarity 31.7%; Pred. No. 1.3e-20;
Matches 77; Conservative 43; Mismatches 104; Indels 19; Gaps 8;

Qy 2 EKSEINEKDLRKSELOGTALGNLKOIYYNSKAITSEKSDAQFLNTLLFKGFFTGH 61
Db 17 ESQDPKPKDELHKSSKFTG-LMENMK-VLYDDNHVSAINKSIDQFLYFDLIYSKDTKL 74

Qy 62 PWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKATACMYGV 111
Db 75 GNYDNRVFEKNKDLADKVKDYVDFGANYYYQCYFSKKTNDINSHQTDKRTKCMYGV 134

Qy 112 TLHDNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKGLYN 171
Db 135 TEHNGNQLDKYRSITVRVFEDEGK-NLLSPD-VQTNKKKVTAEQELDYLTRHYLVKNNKLYE 192

Qy 172 SDSFGKVGKVGRLIVFHSSEGSPVSYDLFDAQGYPD--TLLRIYRDNTTISSTLSLSIY 229
Db 193 FNN--SPIETGYIKFIENENS--FWYDMPAPGDKFDQSKYLMYNDNMKVDKVKIEVY 249
```

```
Qy 112 TLHDNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKGLYN 171
Db 135 TEHNGNQLDKYRSITVRVFEDEGK-NLLSPD-VQTNKKKVTAEQELDYLTRHYLVKNNKLYE 192

Qy 172 SDSFGKVGKVGRLIVFHSSEGSPVSYDLFDAQGYPD--TLLRIYRDNTTISSTLSLSIY 229
Db 193 FNN--SPIETGYIKFIENENS--FWYDMPAPGDKFDQSKYLMYNDNMKVDKVKIEVY 249

Qy 230 LYT 232
Db 250 LYT 252

RESULT 36
US-08-580-806-2
; Sequence 2, Application US/08580806
; Patent No. 5935568
; GENERAL INFORMATION:
; APPLICANT: Dow, Steve W.
; APPLICANT: Elmslie, Robyn E.
; APPLICANT: Potter, Terence A.
; TITLE OF INVENTION: GENE THERAPY FOR EFFECTOR CELL REGULATION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580.806
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2879-29-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-580-806-2

Query Match      22.0%; Score 267.5; DB 2; Length 255;
Best Local Similarity 31.7%; Pred. No. 1.3e-20;
Matches 77; Conservative 43; Mismatches 104; Indels 19; Gaps 8;

Qy 2 EKSEINEKDLRKSELOGTALGNLKOIYYNSKAITSEKSDAQFLNTLLFKGFFTGH 61
Db 17 ESQDPKPKDELHKSSKFTG-LMENMK-VLYDDNHVSAINKSIDQFLYFDLIYSKDTKL 74

Qy 62 PWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKATACMYGV 111
Db 75 GNYDNRVFEKNKDLADKVKDYVDFGANYYYQCYFSKKTNDINSHQTDKRTKCMYGV 134

Qy 112 TLHDNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKGLYN 171
Db 135 TEHNGNQLDKYRSITVRVFEDEGK-NLLSPD-VQTNKKKVTAEQELDYLTRHYLVKNNKLYE 192

Qy 172 SDSFGKVGKVGRLIVFHSSEGSPVSYDLFDAQGYPD--TLLRIYRDNTTISSTLSLSIY 229
Db 193 FNN--SPIETGYIKFIENENS--FWYDMPAPGDKFDQSKYLMYNDNMKVDKVKIEVY 249
```


TELEFAX: (301) 619-7714
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 251
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-144-776B-16

Query Match 21.4%; Score 263.5; DB 4; Length 251;
Best Local Similarity 32.1%; Pred. No. 3.5e-20;
Matches 76; Conservative 45; Mismatches 97; Indels 19; Gaps 10;
Qy 4 SEEINEKDLRKSELOGTALGNLQIYY-YNSKAIT-SSEKSADQFLNTLLFGFTGH 60
Db 25 SOEVAQDPDPQSQRHSSLVKNLQNIYFLYEGDVTTHENKVSVDQLRSHDLIYN---VS 81
Qy 61 HPWYNDLLDLGTAATSEYEGSSVDLYGAYGYQC-AGGTPNKACMYGGVTLDHNNRL 119
Db 82 GPYDKLTKLKNQEMATLFRKDNVDIYGVYHLYCLCENAESACIYGGVTHNEGHL 141
Qy 120 TEKKVPIINLWDGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFGLYNSDFGKV 179
Db 142 EIPKKIVVKSVDIGIQ-SLSFD-IETNKKMVTQAELDYKVRKYLTQDNKQLYTNGP--SKY 197
Qy 180 ORGLIVFHSSEGSTVSVDLFD----AOGYPTDLLRIYRDNTTISSTLSLSISLYLT 232
Db 198 EFGYIKFIPKNKESFWDFPEPEPTQSKY----LMIYKDNELTDSNTSQIEVYLTT 250

RESULT 40
US-08-896-933-26
; Sequence 26, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus aureus

US-08-896-933-26
Query Match 21.4%; Score 260.5; DB 4; Length 239;
Best Local Similarity 32.0%; Pred. No. 6.8e-20;
Matches 77; Conservative 41; Mismatches 104; Indels 19; Gaps 8;
Qy 2 EKSEINEKDLRKSELOGTALGNLQIYYNSKAITSSKSADQFLNTLLFGFTGH 61
Db 1 ESQDPKDELHKSSKFTG-LMENMK-VLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKL 58
Qy 62 PWYNDLLDLGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKACMYGGV 111
Db 59 GNYDNVRVEFNKDLADYKDKYVDVFGANYYQCYFSKKTNDINSHQTDKRTKTCMYGGV 118
Qy 112 TLHDNNRLTEKKVPIINLWDGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFGLYN 171
Db 119 TEHNGNQLDKYRSLTVRVFEDGK-NLLSFD-VQTNKKKVTQAQELDYLTRHYLVKNKKLYE 176
Qy 172 SDSFGGKVQVGLIVFHSSEGSTVSVDLFDAGQOYPD--TLRIYRDNTTISSTLSISLY 229
Db 177 FNN--SPYETGYIKFIENENS-FWYDLMPAPGDKFDQSKYLMNNDNMVDSKDVKIEVY 233

RESULT 41
US-09-314-235-26
; Sequence 26, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; CURRENT FILING DATE: 1999-05-18
; EARLIER FILING DATE: 1997-07-18
; EARLIER FILING DATE: 1994-06-02
; EARLIER FILING DATE: 1992-06-01
; EARLIER FILING DATE: 1991-01-17
; EARLIER FILING DATE: 1990-01-17
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus aureus

Qy 230 L 230
Db 234 L 234
RESULT 41
US-09-314-235-26
; Sequence 26, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; CURRENT FILING DATE: 1999-05-18
; EARLIER FILING DATE: 1997-07-18
; EARLIER FILING DATE: 1994-06-02
; EARLIER FILING DATE: 1992-06-01
; EARLIER FILING DATE: 1991-01-17
; EARLIER FILING DATE: 1990-01-17
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus aureus

US-09-314-235-26
Query Match 21.4%; Score 260.5; DB 4; Length 239;
Best Local Similarity 32.0%; Pred. No. 6.8e-20;
Matches 77; Conservative 41; Mismatches 104; Indels 19; Gaps 8;
Qy 2 EKSEINEKDLRKSELOGTALGNLQIYYNSKAITSSKSADQFLNTLLFGFTGH 61
Db 1 ESQDPKDELHKSSKFTG-LMENMK-VLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKL 58
Qy 62 PWYNDLLDLGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKACMYGGV 111
Db 59 GNYDNVRVEFNKDLADYKDKYVDVFGANYYQCYFSKKTNDINSHQTDKRTKTCMYGGV 118
Qy 112 TLHDNNRLTEKKVPIINLWDGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFGLYN 171
Db 119 TEHNGNQLDKYRSLTVRVFEDGK-NLLSFD-VQTNKKKVTQAQELDYLTRHYLVKNKKLYE 176
Qy 172 SDSFGGKVQVGLIVFHSSEGSTVSVDLFDAGQOYPD--TLRIYRDNTTISSTLSISLY 229
Db 177 FNN--SPYETGYIKFIENENS-FWYDLMPAPGDKFDQSKYLMNNDNMVDSKDVKIEVY 233
Qy 230 L 230
Db 234 L 234

RESULT 42
US-09-414-276-8
; Sequence 8, Application US/09414276
; Patent No. 6392121
; GENERAL INFORMATION:
; APPLICANT: Mason, Hugh
; APPLICANT: Palmer, Kenneth
; APPLICANT: Hefferon, Kathleen
; APPLICANT: Mor. Tsafir
; APPLICANT: Arntzen, Charles
; TITLE OF INVENTION: Gemini Virus Vectors for Gene Expression in Plants
; FILE REFERENCE: 4868/84453


```
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 266
;   TYPE: Amino Acid
;   STRANDEDNESS: Unknown
;   TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-144-776B-8
```

```
Query Match      20.6%; Score 250.5; DB 4; Length 266;
Best Local Similarity 30.7%; Pred. No. 9.4e-19;
Matches 75; Conservative 43; Mismatches 107; Indels 19; Gaps 8;

QY 1 SEKSEINEKDLRKKSELOQTALGNLQIYYVNSKAITSEKSDAQFLNTLLFKGFFTG 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 27 AESQDPKPDELHKSSKFTG-LMENNK-VLYDDNHVSAINVKSIDQFRYFDLIYSIKDTK 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 HPWYNDLLVGLGSTAATSEYEGSSVDLYGAYGYQCAG-----TPNKTACMYGG 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 LGYDNRVVEFKNKDLADKYDKYVDVEGANAYYQCAFSSKKTNDINSHQTDKRTKCMYGG 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 111 VTLDHNNRLTEKKVPINLWIDGKQTVPIDKVKTSKEVTQVQLDQARHYLHGKFGCL 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 145 VTEHGNQLDKYRSITVRVFDGK-NLLSFD-VQYNNKKVTAQELDYLTRHYLVKNKKLY 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 171 NSDSFGGVQVRGLIVFHSSEGSTVSYDLFDAQGOYPD--TLIRYRDNNTTISSTLSLS 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 203 EPNN--SPYETGYIKFIENENS-FWYDMPAPGDKFDQSKYLMYNDKNWYDSKDVKIEV 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 229 YLYT 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 YLYT 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 47
US-08-896-933-21
; Sequence 21, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/003002
; CURRENT APPLICATION NUMBER: US/08/896.933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-21
```

```
Query Match      18.5%; Score 225.5; DB 4; Length 239;
Best Local Similarity 30.2%; Pred. No. 3.8e-16;
Matches 74; Conservative 43; Mismatches 105; Indels 23; Gaps 11;

QY 2 EKSEINEKDLRKKSELOQTALGNLQIYYVNSKAITSEKSDAQFLNTLLFKGFFTG 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 ESQDPKPDELHKSSKFTG-LMENNK-VLYNDHVSAINVKSINEFFDLIYLSIKDTKL 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 PWYNDLLVGLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGG 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 GNYDNRVVEFKNKDLADKYDKYVDVEGANAY-YQCYFSKKTNNIDSHENTKRKT-CMYG 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 111 VTLDHNNRLTE-EKKVPINLWIDGKQTVPIDKVKTSKEVTQVQLDQARHYLHGKFGCL 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 VTEHGNQLDKYRSITVRVFDGK-NLLSFD-VQYNNKKVTAQELDYLTRHYLVKNKKL 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 170 YNSDSFGGVQVRGLIVFHSSEGSTVSYDLFDAQGOYPD--TLIRYRDNNTTISSTLSLS 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 229 YLYT 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 YLYT 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 175 YEFNN--SPYETGYIKFIENENS-FWYDMPAPGNKFDQSKYLMYNDKNWYDSKDVKIE 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 228 YLYT 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 232 YLYT 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 48
US-09-314-235-21
; Sequence 21, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; EARLIER FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416,530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-21
```

```
Query Match      18.5%; Score 225.5; DB 4; Length 239;
Best Local Similarity 30.2%; Pred. No. 3.8e-16;
Matches 74; Conservative 43; Mismatches 105; Indels 23; Gaps 11;

QY 2 EKSEINEKDLRKKSELOQTALGNLQIYYVNSKAITSEKSDAQFLNTLLFKGFFTG 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 ESQDPKPDELHKSSKFTG-LMENNK-VLYNDHVSAINVKSINEFFDLIYLSIKDTKL 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 PWYNDLLVGLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGG 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 GNYDNRVVEFKNKDLADKYDKYVDVEGANAY-YQCYFSKKTNNIDSHENTKRKT-CMYG 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 111 VTLDHNNRLTE-EKKVPINLWIDGKQTVPIDKVKTSKEVTQVQLDQARHYLHGKFGCL 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 VTEHGNQLDKYRSITVRVFDGK-NLLSFD-VQYNNKKVTAQELDYLTRHYLVKNKKL 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 170 YNSDSFGGVQVRGLIVFHSSEGSTVSYDLFDAQGOYPD--TLIRYRDNNTTISSTLSLS 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 YEFNN--SPYETGYIKFIENENS-FWYDMPAPGNKFDQSKYLMYNDKNWYDSKDVKIE 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 228 YLYT 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 232 YLYT 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 49
US-08-896-933-20
; Sequence 20, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005002
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:58:32 ; Search time 4.3768 Seconds
(without alignments)
1032.821 Million cell updates/sec

Title: US-09-900-766-2

Perfect score: 1218

Sequence: 1 SEKSEINEKDLRKKELOG.....RDNTTISLSISLYLYTT 233

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1072	88.0	248	9	US-09-870-759-16
2	935	76.8	257	9	US-09-870-759-8
3	611	50.2	258	9	US-09-870-759-14
4	268.5	22.0	251	8	US-08-973-391A-13
5	268.5	22.0	266	9	US-09-870-759-10
6	267.5	22.0	239	10	US-09-150-947B-12
7	263.5	21.6	266	9	US-09-870-759-12
8	220	18.1	250	9	US-09-870-759-20
9	100	8.2	234	9	US-09-870-759-18
10	87	7.1	226	10	US-09-815-242-5900
11	87	7.1	226	10	US-09-815-242-13156
12	82	6.7	772	9	US-10-121-032-28
13	79	6.5	833	10	US-09-844-281-1
14	78	6.4	607	10	US-09-815-242-13379
15	78	6.4	607	10	US-09-815-242-13682
16	78	6.4	1551	10	US-09-864-761-35904
17	77.5	6.4	312	10	US-09-815-242-4904
18	77.5	6.4	312	10	US-09-815-242-10541
19	77.5	6.4	985	9	US-09-738-626-4377

ALIGNMENTS

RESULT 1
US-09-870-759-16
; Sequence 16, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-870-759-16

Query Match 88.0%; Score 1072; DB 9; Length 248;
Best Local Similarity 90.2%; Pred. No. 2.1e-100;
Matches 202; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Qy	1	SEKSEINEKDLRKKELOGTALGNLQIYYNSKAITTSEKSADQFLNTLLFKGFFTG	60
Db	25	SEKSEINEKDLRKKELOGTALGNLQIYYNEKAITENKESDDQFLNTLLFKGFFTG	84
Qy	61	HPWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHDNNRLT	120
Db	85	HPWYNLLVDLGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTCMYGGVTLHDNNRLT	144
Qy	121	EKKVPINLWIDGKQTVTDIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDFGKQV	180
Db	145	EKKVPINLWIDGKQTVTDIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDFGKQV	204
Qy	181	RGLIVFHSSEGSTVSYDLFDAQGYPTDLLRIYRONTTISSTSL	224
Db	205	RGLIVFHSSEGSTVSYDLFDAQGYPTDLLRIYRONKTNSENL	248

RESULT 2

```
US-09-870-759-8
; Sequence 8, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 257
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-870-759-8

Query Match 76.8%; Score 935; DB 9; Length 257;
Best Local Similarity 75.5%; Pred. No. 1.3e-86;
Matches 176; Conservative 21; Mismatches 36; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKKSELOCTALGNLKOIYYNNSKAITSEKSADQFLNTLLFKGFFTG 60
Db 25 SEKSEINEKDLRKKSELOCTALGNLKOIYYNNEKAKTENKESHQFLQHTILFKGFFTN 84
QY 61 HPWYNLLVDLGSTAASTSEYSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 85 HSNYDNLVDFDSKIDVYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFGGKVQ 180
Db 145 EEKKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFGGKVQ 204
QY 181 RGLIVPHSSEGSTSVYDLFDAGQYPTDLLRIYRDNTTISSTLSISLYLT 233
Db 205 RGLIVPHSTEPSVYDLFDAGQYPTDLLRIYRDNTTISSTLSISLYLT 257

RESULT 3
US-09-870-759-14
; Sequence 14, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 258
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-870-759-14

Query Match 50.2%; Score 611; DB 9; Length 258;
Best Local Similarity 51.1%; Pred. No. 5.3e-54;
Matches 118; Conservative 35; Mismatches 78; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKKSELOCTALGNLKOIYYNNSKAITSEKSADQFLNTLLFKGFFTG 60
Db 26 NENIDSVKEKELHKKSELSSTALNNMKHSYADKNPIGENKSTGDFLENTLLYKFFTD 85
QY 61 HPWYNLLVDLGSTAASTSEYSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 86 LNFEDLLINFNSKEMAQHFKNVDVYPIRISYNGYGEIDRTACTYGGVTPHEGNK 145
QY 121 EEKKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFGGKVQ 180

US-09-870-759-10
; Sequence 10, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 266
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-870-759-10

Query Match 22.0%; Score 268.5; DB 8; Length 251;
Best Local Similarity 32.5%; Pred. No. 1.4e-19;
Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SEEINEKDLRKKSELOCTALGNLKOIYYNNSKAITSEKSADQFLNTLLFKGFFTG 60
Db 25 SOEVAQADPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENKSVQDOLLSHHLYN---VS 81
QY 61 HPWYNLLVDLGSTAASTSEYSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRL 119
Db 82 GPNYDKLKTENKQEMATLFKDKNDVIYGVYVHLCYLCENASACIYGGVTNHEGNHL 141
QY 120 TEKKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFGGKV 179
Db 142 EIPKIVVKSIDGIGQ-SLSFD-IETNKKMVTQAQLDYKRVKYLTDNKKOLYTNGP--SKY 197
QY 180 QRGLIVPHSSEGSTSVYDLFD---AQGYPTDLLRIYRDNTTISSTLSISLYLT 232
Db 198 ETGYIKFIPKNKESFWDFPEPEFTQSKY----LMIYKDNETLDSNTSQIEVILTT 250

RESULT 5
US-09-870-759-10
; Sequence 10, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 266
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-870-759-10

US-09-900-766-2.2.rapb
; Sequence 13, Application US/09973391A
; Patent No. US20020054887A1
; GENERAL INFORMATION:
; APPLICANT: Schlievert, Patrick M.
; APPLICANT: Roggiani, Manuela
; APPLICANT: Stoehr, Jennifer
; APPLICANT: Ohlendorf, Douglas
; TITLE OF INVENTION: MUTANTS OF STREPTOCOCCAL TOXIN A AND METHODS OF USE
; FILE REFERENCE: 600.311USWO
; CURRENT APPLICATION NUMBER: US/08/973,391A
; CURRENT FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: PCT/US96/10252
; PRIOR FILING DATE: 1996-06-07
; PRIOR APPLICATION NUMBER: US 08/480,261
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 251
; TYPE: PRF
; ORGANISM: Streptococcus pyogenes
US-09-973-391A-13

Db 146 ERKKIPINLWINGQKEVSLDKVQTDKNVTYQELDAQARRYLQDKLKLNNDTLGGKIQ 205
QY 181 RGLIVPHSSEGSTSVYDLFDAGQYPTDLLRIYRDNTTISSTLSISLYLY 231
Db 206 RGKIEFDSSDGSKSVYDLFDVKGDFPEKOLRIYSDNKTLSLSTELHDIYLY 256

RESULT 4
US-08-973-391A-13
; Sequence 13, Application US/09973391A
; Patent No. US20020054887A1
; GENERAL INFORMATION:
; APPLICANT: Schlievert, Patrick M.
; APPLICANT: Roggiani, Manuela
; APPLICANT: Stoehr, Jennifer
; APPLICANT: Ohlendorf, Douglas
; TITLE OF INVENTION: MUTANTS OF STREPTOCOCCAL TOXIN A AND METHODS OF USE
; FILE REFERENCE: 600.311USWO
; CURRENT APPLICATION NUMBER: US/08/973,391A
; CURRENT FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: PCT/US96/10252
; PRIOR FILING DATE: 1996-06-07
; PRIOR APPLICATION NUMBER: US 08/480,261
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 251
; TYPE: PRF
; ORGANISM: Streptococcus pyogenes
US-08-973-391A-13

Db 146 ERKKIPINLWINGQKEVSLDKVQTDKNVTYQELDAQARRYLQDKLKLNNDTLGGKIQ 205
QY 181 RGLIVPHSSEGSTSVYDLFDAGQYPTDLLRIYRDNTTISSTLSISLYLY 231
Db 206 RGKIEFDSSDGSKSVYDLFDVKGDFPEKOLRIYSDNKTLSLSTELHDIYLY 256
```



```
Db 35 KPSQLORSNLVTKFIYIFMRVTLVTHENVKSVLDQLSHDLIYN---VSGPNYDKLTKTE 91
QY 71 LGSTAATSEYSGSSVDLYGAYGYQC-AGTGNKTCACMYGGVTLHDNNRLTEKKVPINL 129
Db 92 LKQEMATLFDKKNVDIYGVYHCLYGCENASACLYGGVTNHGHNLEIPKIKIVKV 151
QY 130 WIDGKOT-TVPIDKVTSKKEVTQVQLDQARHYLHGKFGLYNSDSFGKGQVGRGLIVFHS 188
Db 152 SIDGIOSLFDIEQING-----NCSRISTVVRKYLTDNKLQTYNGP--SKYETGYIKFIP 205
QY 189 SEGTSVSYDLFD-----AQOQYPTDLRIYRDNTTISSTLSISLYLT 232
Db 206 KKNESWFDFFPEPTQSKY-----LMIVKNETLDSNTSQIEVLYLT 249

RESULT 9
US-09-870-759-18
; Sequence 18, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 18
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-870-759-18

Query Match 8.2%; Score 100; DB 9; Length 234;
Best Local Similarity 24.5%; Pred. No. 0.012;
Matches 48; Conservative 33; Mismatches 61; Indels 54; Gaps 12;

QY 25 NLKQI---YYNSKATTSSE-----KSADQFLNTLLFKGFFTGHPWYNDLL-- 58
Db 45 NIKDLLDWSYSSGSDTETNSEVLNDSLGSMRIKNTDGSI-SLIIFPS-----PYISPAFTK 98
QY 69 ---VDLGGSTAATSEYSGSSVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLTEKKV 125
Db 99 GEKVDL-NTRKTKSHTS---EGTIIHFQISGVT-----NTEKLPTPIEL 140
QY 126 PINLWIDGKQTTVPIDKVTSKKEVTQVQLDQARHYLHGKFGLY-NSDSFGG--KVQRG 182
Db 141 PLKVKVHGKDSPLKYGP-RFDKKQLAISTLDPEIRHQLTQIHGLYRSSDKTGGYWKIT-- 197
QY 183 LIVFHSSEGSTVSVDL 198
Db 198 -----MNDGSTVQSGL 208

RESULT 10
US-09-815-242-5900
; Sequence 5900, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5900
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5900

Query Match 7.1%; Score 87; DB 10; Length 226;
Best Local Similarity 21.9%; Pred. No. 0.23;
Matches 42; Conservative 42; Mismatches 92; Indels 16; Gaps 6;

QY 8 NEKDRLKRSLOGTALGNLKIYYNSKAITSEKSADQFLNTLLFKGFFTGHPWYNDL 67
Db 24 NVQSVOAKAEVQKQSESELK--HYTNKPILEKRNVTGFKYTDGKHYLEVTVGQ--QHSR 79
QY 68 LVDLGSTA-ATSEYSGSSVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLTEKKVP 126
Db 80 ITLLGSKDKFKDGENSNIDVF-----ILREGDSROATNYSIGGVTKSNVQYIDINTP 134
QY 127 INLWIDGKQTTVPIDKVTSKKEVTQVQLDQARHYLHGKFGLYNSDSFGGKVQRGLIVF 186
Db 135 I-LEIKKQNEVDLKFQYYSKEDISLKELDYRLRERAIKQHGLYSNG-----LKQGOITI 188
QY 187 HSSEGSTVSVDL 198
Db 189 TWDGTTHTIDL 200

RESULT 11
US-09-815-242-13156
; Sequence 13156, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; CURRENT APPLICATION NUMBER: 60/257,931
```

07 232 11: 623 YTS 625
 Db
 RESULT 14
 US-09-815-242-13379
 ; Sequence 13379, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert


```
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35904
; LENGTH: 1551
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL050341.18
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EST_HUMAN HIT: AV715688.1, EVALUE 1.00e-129
; OTHER INFORMATION: SWISSPROT HIT: Q06730, EVALUE 5.00e-15
US-09-864-761-35904

Query Match 6.4%; Score 78; DB 10; Length 1551;
Best Local Similarity 23.9%; Pred. No. 26;
Matches 51; Conservative 31; Mismatches 55; Indels 76; Gaps 11;

QY 7 INEKDLRKSELQGTALGNLKIYYNS--KAITSEKSADQF-----LNTLTLF 54
Db 1382 VPKENFRKHQSPRDLKTYKPMGFESSFLKFIQSEEEKEDFDWEPSEHITLSNSSOS 1441
QY 55 KGFFTGHPWYNLLVDLGSTAATSEYSGVDLYGAYGYQCAGGTPNKTACMYGGVTLH 114
Db 1442 SNDLTGNVYANNVND-----SEPE---VDI-----PHSSS-----DSTTH 1474
QY 115 DNNRLTEEEKVPINLWIDGKQTTVP-----IDKVKTSKKEVTVQVQLDLOARHYLHCK 166
Db 1475 EN--LT-----ALPPLIVAETTTVPFSLNLRVLDKALDCGELAKQL-----HYL---- 1519
QY 167 FGLYNSDSFGKVGKQGLIVFHSSEGSTVSYDLF 199
Db 1520 -----RPVVVLERSKESFTPLDLF 1538

RESULT 17
US-09-815-242-4904
; Sequence 4904, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10541
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FILE REFERENCE: ELITRA.011A
; TITLE OF INVENTION: Identification of Essential Genes in
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4904
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-4904

Query Match 6.4%; Score 77.5; DB 10; Length 312;
Best Local Similarity 34.0%; Pred. No. 3.2;
Matches 16; Conservative 12; Mismatches 16; Indels 3; Gaps 1;

QY 119 LTRKKVPINLWIDGKQTTVPIDKVKTSKK---EVTVQVQLDLOARHY 162
Db 215 MTEHLDFVNLDEAKNAEIMVEKETQKMLETFEELDLSVRSY 261

RESULT 18
US-09-815-242-10541
; Sequence 10541, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10541
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10541

Query Match 6.4%; Score 77.5; DB 10; Length 312;
Best Local Similarity 34.0%; Pred. No. 3.2;
```



```
; SEQ ID NO 215
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-215

Query Match      6.3%; Score 77; DB 9; Length 579;
Best Local Similarity 24.1%; Pred. No. 8.4;
Matches 33; Conservative 13; Mismatches 51; Indels 40; Gaps 3;

QY 71 LGSTAATSEYEGSSVDLYGAYGYOCAGTPNKTACM-----YGVVTLH 114
| : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 306 VGGNAAGVGGSHGLVGSNSIGGVNSVPNSNAMVGGGLGSGSSGSGAGGEHLN 365
| : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY 115 DNNRLTEKKVPINLWIDGKOTTP-----IDKVKTSKKEVTQ----- 153
| : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 366 DNSSNDKLKSGVQVSPDGKVTNIPATVWNNQFMVGLLTFIRAAETDPLVTLGLGDL 425
| : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY 154 ---ELDQARHYLHGKF 167
| : : : | : : | : : |
Db 426 TGLGLNLNSQESLHTTF 442
| : : : | : : | : : |

RESULT 22
US-09-801-368-218
; Sequence 218, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 218
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-218

Query Match      6.3%; Score 77; DB 10; Length 704;
Best Local Similarity 21.5%; Pred. No. 11;
Matches 42; Conservative 29; Mismatches 76; Indels 48; Gaps 7;

QY 38 TSSEKSAQOFLNTLLFKGFFTHGPWYNDLGLGTAATSEYEGSSVDLYGAYGYOCA 97
| : : : | : : | : : | : : : | : : : | : : : | : : : |
Db 333 TSSSNASPNTTMMANADSNAGNPKNDATIDNELTOILNEYNNFNDNLGT-----S 386
| : : : | : : | : : | : : : | : : : | : : : | : : : |
QY 98 GGTPTNKTACMGVTLHNNRLT-----EEKKVPINLWIDGKOTTPIDKVKTSKKEVTY 152
| : : : | : : | : : | : : : | : : : | : : : | : : : |
Db 387 TSGKNKSAC-----PSSFOANAMTKINPSOOLQQLNR-VQHKQLTSSHNNSSTNMKS--- 438
| : : : | : : | : : | : : : | : : : | : : : | : : : |
QY 153 QELDQARHYLHGKGLGYNDSFGKVGKVGRLIVFHSSSEGSTVSYDLFDAQGYQPTLLRI 212
| : : : | : : | : : | : : : | : : : | : : : | : : : |
Db 439 -----FNSDLSYRRQRASLPIDDS-----LSYDLVNKQDEDP----- 471
| : : : | : : | : : | : : : | : : : | : : : | : : : |
```

```
QY 213 YRDNTTISSTSLIS 227
| : : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 472 --KNDMLPNSNLSS 484

RESULT 23
US-09-815-242-13445
; Sequence 13445, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13445
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13445

Query Match      6.3%; Score 76.5; DB 10; Length 454;
Best Local Similarity 18.3%; Pred. No. 6.8;
Matches 39; Conservative 32; Mismatches 73; Indels 69; Gaps 8;

QY 5 BEINEKDLRKKSSELOG-----TALGNL---KOIYYNSKAITSSEKSADQ 46
| : : : | : : | : : | : : : | : : : | : : : | : : : |
Db 20 EDFSDGVRAAEEGFEALCGWEASEASOLLNLPPELHYQNNSSELANGKEK----- 75
| : : : | : : | : : | : : : | : : : | : : : | : : : |
QY 47 FLTNTLLFKGFFTHGP-----WYNLLVDLGSTAAT----- 77
| : : : | : : | : : | : : : | : : : | : : : | : : : |
Db 76 --VKVLLAKALF-GKPDVLLDEPTNGLDIQSITWLEDFLIDFNTVIWVSHDRHFNKV 132
| : : : | : : | : : | : : : | : : : | : : : | : : : |
QY 78 ----SEYEGSSVDLYGAYGYOCAGTPNKTACMTGGVTLHNNRLTEKKVPINLWIDG 133
| : : : | : : | : : | : : : | : : : | : : : | : : : |
Db 133 CTHMADLDFGKTKLVGNVDFW-----KESSELAALLADRNAKAEKIKQLQEEFVAR 185
| : : : | : : | : : | : : : | : : : | : : : | : : : |
QY 134 KQTTVPIDKVKTSK-----EVTVQELDLQARHY 162
| : : : | : : | : : | : : : | : : : | : : : | : : : |
Db 186 FSNASKSRQATSRKKMLDKIELEIVPSSRY 218
| : : : | : : | : : | : : : | : : : | : : : | : : : |

RESULT 24
US-09-738-626-4306
; Sequence 4306, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
```

```

; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4306
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4306

Query Match          6.2%  Score 75.5; DB 9; Length 424;
Best Local Similarity 20.8%; Pred. No. 7.8;
Matches 38; Conservative 26; Mismatches 56; Indels 53; Gaps 9;

Qy 74 TAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMGVTLHDNNRLT---EKKVPIPLNW 130
Db 196 TTQLKQVEGLSVNTIGFEG-----WGSVLDDGNGVTVDSDAKAGLQAL 241

Qy 131 IDG-KQTTVPIDKVKTSKEVTVOELDLQARH-----YLFH-----GKF---GLYN 171
Db 242 VDGDDCTISKASLAATEENLAFTEGQTAYAINWPMYTNSEAEATAGKFEVQPLVG 301

Qy 172 SDSFG---GRVQGLIVFHSSEGSTVSVDLFD-----AQQGYPTDLLRIYRDN 216
Db 302 KDGVGSTLGGYNGINV--NSENKATARDPIEFIEENQTFADNSFPVPLASIYDDE 359

Qy 217 TTI 219
Db 360 SLV 362

RESULT 25
US-10-055-364-44
; Sequence 44, Application US/10055364
; Patent No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Equine herpesvirus 5
; US-10-055-364-44

Query Match          6.2%  Score 75.5; DB 9; Length 818;
Best Local Similarity 22.3%; Pred. No. 19;

```

```

Matches 52; Conservative 33; Mismatches 81; Indels 67; Gaps 13;

Qy 6 EINEKDLRKKSELOQTALGNLKOIYYVNSKAITSSSEKSAQDLFTNTLLFKGFFTGHPWYN 65
Db 73 EENCPTNTEDEHVEGILL-----IYKTNIVPYIFNVKRYKRLVSTTIYKG-----WSQ 121

Qy 66 DLLVDLGLSTAATSEYEGS-SVDLYGAY---YGYQCAGGTPNKTACMGVTLHDNNRLT 121
Db 122 D-----AITNQYTSFAMPLWEARLVDVNYEC-----YNGIQVTENGHLT- 161

Qy 122 EKKVPIPLNWIDGKQTTVPIDKVKTSKEVTVOELDLQARHLYHGKFGLYNSDSFGGKQVOR 181
Db 162 -----TYVDRDGYNESVRLVPADGLTSSIRRY-HSQPELY-----VTPR 199

Qy 182 GLIVFHSSEGSTVSVDLFD--AQQGYPTDLLRIYRDNTTISSTLSISLVLTYT 232
Db 200 NLLWSYTR-TTVNCEVIDMTARSHKP-----FEYFVTASGDSIETSPF-YT 244

RESULT 26
US-09-738-626-4498
; Sequence 4498, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4498
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4498

Query Match          6.1%  Score 74.5; DB 9; Length 386;
Best Local Similarity 21.4%; Pred. No. 8.6;
Matches 41; Conservative 23; Mismatches 67; Indels 61; Gaps 8;

Qy 23 LGNLKQIYYVNSKA-ITSSEKSAQDLFTNTLLFKGFFTGHP-----WYNDLILD 70
Db 131 LVNSALVYYSSNAFLHAAEKLGTSSFSHTFVSGEVTGIDPEELEDGCGFTFYERDLAE 190

Qy 71 LG-----STAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMY-----GGVT-- 112
Db 191 LGPHFDRLPEDMKAFALNDGEGDRLEQFEQINSFVC---RPRHTGNQFLAMQTKGAKTPY 247

Qy 113 -----LHDNNRLTEKKVPIPLNWIDGKQTTVPIDKVKTSKEVTVOELDLQARHLYHG 165
Db 248 IPRHFHRLHTENFIC-----LDGR-----VKLHVNGQEITLSRGDYVHA 286

Qy 166 KFGLYNSDSFGG 177
Db 287 PAGTIHSFAFAG 298

```

RESULT 27

US-09-815-242-10709
; Sequence 10709, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10709

; LENGTH: 667

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-815-242-10709

Query Match 6.1%; Score 74.5; DB 10; Length 667;

Best Local Similarity 17.8%; Pred. No. 18;

Matches 63; Conservative 36; Mismatches 86; Indels 169; Gaps 14;

QY 16 SELOGTALCN-----LKOIYYNSKAITSSSEKSADQFLTNL-----LFLKGF 58

DB 263 SSVHGAPICEEITAAKAVGHEYPDFTVPVEVAARFKETMIDEGQKAEANMEFKNVE 322

QY 59 TGHF-----WYNDL-LVDLGSTAA----- 76

DB 323 HARPELAKQKFAFANQLPEGNEOELPKYELGTSAASRVTSKETIQASIKVVPSEWGGSA 382

QY 77 -----TSEYEGSSV-----DLYGAYYGQCAGGTPNKTACHMYGGV- 111

DB 383 DLSASNTMTAAEKOFESQYEGRNWFGVREFAMAAMNGTLHGSSH-----VYGGTF 437

QY 112 -----TLHNNRLTEKKVPI-----NLWT----- 131

DB 438 FVFTDYLRPAIRLAALOKPVTVYVTHDSVAVGEDGTPHEIQLASVRCIPNVHVIRPA 497

QY 132 DGKQTTVPIDKVKTSKKEVTQVQELQARHYLHGKFGLYNSDSFGKVGORGLIVFHSSEG 191

DB 498 DGNETVAANKIAMTSTETITLVLSRQNLPLVLEGTLE-HASDS-----VQKAYVLSPQKG 552

QY 192 -----STVSVDLFDQAQ-GQYPTDLL 210

DB 553 EQPSGILITATSGEVNLAVEAQAALAEEDIDVSVWSPSPDLFEKQSAEYKESVL 606

RESULT 28

US-09-815-242-13405
; Sequence 13405, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13405

; LENGTH: 750

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-815-242-13405

Query Match 6.1%; Score 74; DB 10; Length 750;

Best Local Similarity 22.0%; Pred. No. 24;

Matches 53; Conservative 34; Mismatches 90; Indels 64; Gaps 11;

QY 26 LKOIYYNSKAITSSSEKSADQFLTN-----TLFLKGFFTGHGPHY-----NDLLVDLGST 74

DB 328 LAQIVDFVKKAQTS--RAPIDBLTKISGIFVPVVVILGIMTFWVFLRLDSVVVLGAS 385

QY 75 AATSEYEGSSVDLYGAYYGQCAGGTPNKTACHMYG-----GVTLDHNNRLTEKKVPI 127

DB 386 FVSSLLYGAVVLIAC----PCALGLATPTALMGVGRSAKMGVLLKNGTVLQETOKVQT 441

QY 128 NLW-----IDGKQTTVPI--DKVRTSKKEVTQVE-----LDLQA-- 159

DB 442 LVFDKGTGLTEGKPVVVDIIGDEVEVFGLAASLEDAASQHPLAEAIVKRASEAGLEFQTV 501

QY 160 -RHYLHGKFCGLNSDSFGKVGORGLIVFHSSEGSTVSYDLFDAQGOYPTDLLRIYRDNTT 218

DB 502 NFOALHGK-----GVSGRINGKQVLL-----GNAKMLDGMDSINTYQDKLELEKEAKT 550

QY 219 I 219

DB 551 V 551

RESULT 29

US-09-815-242-10668
; Sequence 10668, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

```

; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10668
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10668

Query Match          6.0%; Score 73.5; DB 10; Length 882;
Best Local Similarity 19.1%; Pred. No. 34;
Matches 50; Conservative 41; Mismatches 112; Indels 59; Gaps 9;

Qy 16 SELQGTALGNLKOIYYNSKAITSSEKSADQFLNTL-----LFGGFTG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 335 SPIVGLANGNNKKMYTNNLAVFESQFIDWLMDETRKKVYDAKRYVALNRVYGMWG 394
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 HPW-----YNDLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGG 110
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 395 IAFDVLLAAVLLDNDNADIEGVAHQYDAIOSD--EAIYKGAKKGLPEDEEVFFGH 452
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 111 -----VTLHDNNRLTEEK-----KVPINLWI-DGKOTTPIDKVKTSKKEVT 151
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 453 LARKIKAIQFLTSLKDLSELTEKNOADLYFKMELPLSRILGDMETITGIRVDATRLKEMQVE 512
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 152 VQELDQARHYLHGKFGLYNSDSFGGKVQVORGLIVFH-----SSEGSVSVYDLFDAQ 202
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 513 FSERLKIEEKIYAEAG--EEFNLSPKQLGVLIFKMGPLVIAKKTKGTSTAVDVLEQL 570
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 203 GQYPTDLLRI--YRDNTTISST 222
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 571 KEQAPIVADILTYRQIAKIQST 592
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 30
US-09-815-242-4994
; Sequence 4994, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10626
; LENGTH: 1207

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4994
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-4994

Query Match          6.0%; Score 73.5; DB 10; Length 1158;
Best Local Similarity 24.4%; Pred. No. 49;
Matches 31; Conservative 16; Mismatches 65; Indels 15; Gaps 2;

Qy 111 VTLHDNNRLTEKKVPINLWIDGKOTTPIDKVKTSKKEVTVQELDQARHYLHGKFGLY 170
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 94 VTLRLTNRETGEIKSQEVFFGDFPLMTMGTFIINGAERVIVSQLVSPGVYFHGKVDKN 153
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 171 NSDSFGGKV--ORGLIVPHSSEGSTVSYDLFDAQOQYP-----DTLLRIYRD 215
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 154 GKEGFGSTVIPNRGAWLEMETDAKDISYVRIDRTKRKIPLTLVLRALGFGSDOTIFEIFGD 213
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 216 NTTISST 222
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 214 SESLRNT 220
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 31
US-09-815-242-10626
; Sequence 10626, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10626
; LENGTH: 1207
```

TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10626

Query Match 6.0%; Score 73.5; DB 10; Length 1207;
Best Local Similarity 24.4%; Pred. No. 52;
Matches 31; Conservative 16; Mismatches 65; Indels 15; Gaps 2;
QY 111 VTLHNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLY 170
Db 97 VTLRLTNRETGIKSOEVFFGDFPLMTENGTEFLNGAERVIVSQLVRSGVYFHGKVDKN 156
QY 171 NSDSFGKV-ORGLIVFHSSEGSTVSVDLFDAGQGYP-----DTLRIYRD 215
Db 157 GKEGFGSTVIPNRGAWLEMETDAKDISYVRIDTRKIPLTVLVRLGFGSDDTIFEIGD 216
QY 216 NNTIST 222
Db 217 SESLNT 223

RESULT 32
US-09-054-141-2
Sequence 2, Application US/09054141
Patent No. US20010018193A1
GENERAL INFORMATION:
APPLICANT: DOUGAN, Gordon
APPLICANT: FRANKEL, Gad
TITLE OF INVENTION: Antibodies to Intimin-like
TITLE OF INVENTION: Proteins of E. coli
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE and DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,141
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,452
FILING DATE: 23-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Hollie L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 102286.319
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 275 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

Query Match 6.0%; Score 73; DB 10; Length 275;
Best Local Similarity 21.4%; Pred. No. 7.6;
Matches 47; Conservative 32; Mismatches 77; Indels 64; Gaps 11;
QY 21 TALGNLQIYYNSKAITSSEKSAOQTLNTLLFKGFFTGHPWYNDLLVDLGLSTAATSY 80
Db 10 TAVANGKAIKYTVKMNQPVNQSFTSTNF-GMFNGK-----SQTQATTGN 58

QY 81 EG-SSVDLYGAYYGQACGTPNKTCMYG-GVTLHNNRLTEKKVPI-----N 128
Db 59 DGRATITLSSAGKATVSATVSDGAEVKATEVTFDEDKI--DNKVKIIGNVRGELPN 116
QY 129 LWDGKQTTVPIDKVKTS-----KKEVTVOELDLQARHYLHGKFGLYNSDSFGKV 179
Db 117 IWLQYGF-----KLKASGGDGTYSWYSENTSIATVDASGKVTLNGK----- 158
QY 180 ORGLIVFHSSEG--STVSVDL-----FDAQGYPTDL 209
Db 159 --GSVVIKATSGDKQTVSYTIKAPSYMIVKIVKQAYYADAM 196
RESULT 33
US-09-740-452-2
Sequence 2, Application US/09740452
Patent No. US20010010912A1
GENERAL INFORMATION:
APPLICANT: Wilding, Edwina Imogen
APPLICANT: Black, Michael T.
APPLICANT: Shilling, Lisa K.
APPLICANT: Kosmatka, Anna L.
APPLICANT: Jaworski, Deborah D.
TITLE OF INVENTION: nrdF
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Fast-SEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/740,452
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/092,437
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10155
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-740-452-2

Query Match 6.0%; Score 73; DB 10; Length 320;
Best Local Similarity 25.2%; Pred. No. 9.4;
Matches 31; Conservative 18; Mismatches 40; Indels 34; Gaps 6;
QY 12 LRKKSLOCTALGNLKOIYYNSKAITSSEKSAOQTLNTLLFKGFFTGHPWY---NDLL 68
Db 130 LOKKAEI-----VNEIYLNQSP---EKKVASVLETFELFYSGFFT--PLYLGNKL 177
QY 69 VDLGSTAATSEYSGSVOLYGAYYGQACGTPNKTCMYGTVTLHNNRLTEKKVPI 128
Db 178 ANVAEIKLIIRDES---VHGTYIGYKFGQGF-----NLPPEEQEKLK 218

[illegible]


```
Qy 124 KVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHLYHGKFGLYNSDSFGGKVQRGL 183
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 240 SNTIE-----TTTAKLEKSRGAKK-RIKERALMGs--LHSTL---NGNSIAGSIQ--- 285

Qy 184 IVFHSSEGSTVSYDLFD---AQGQY---PDTLLRIRYDNTTISSTLSIS 227
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 286 -----TISHDLYDDDSMOGAFDNPSSFRPRTQSNLSIPGSSSRVS 326

RESULT 40
US-09-844-353A-45
; Sequence 45, Application US/09844353A
; Patent No. US20020037585A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowsek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; CURRENT APPLICATION NUMBER: US/09/844,353A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 08/857,076
; PRIOR FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-844-353A-45

Query Match 6.0%; Score 73; DB 10; Length 510;
Best Local Similarity 20.9%; Pred. No. 18;
Matches 48; Conservative 37; Mismatches 11; Indels 34; Gaps 9;

Qy 4 SBEINEKDLRKSELQGTALGNLKOIYYNSKAITSSSEKADQFLTNTLLFKGFTGHPW 63
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 125 SESPDVTSGKKTTTRNAWGNMVAELITTAIMASPEKR---LTLAQVYEMWQVNPY 180

Qy 64 YNDLLVLGSTAATSEYEGSSVDLYGAYGQCAGGTPNKACMYGGVTLHDNNRLTEEK 123
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 181 FRD-KGDSNSSAGMKNSIRHNLSLHSRFMRIQNEGAGKSSWWVINPDAPGMPNPRTRER 239

Qy 124 KVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHLYHGKFGLYNSDSFGGKVQRGL 183
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 240 SNTIE-----TTTAKLEKSRGAKK-RIKERALMGs--LHSTL---NGNSIAGSIQ--- 285

Qy 184 IVFHSSEGSTVSYDLFD---AQGQY---PDTLLRIRYDNTTISSTLSIS 227
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 286 -----TISHDLYDDDSMOGAFDNPSSFRPRTQSNLSIPGSSSRVS 326

RESULT 41
US-09-801-368-418
; Sequence 418, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Royer, John
```

```
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 418
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-418

Query Match 6.0%; Score 73; DB 10; Length 566;
Best Local Similarity 20.5%; Pred. No. 21;
Matches 38; Conservative 31; Mismatches 50; Indels 66; Gaps 8;

Qy 7 INEKDLRKSELQGTALGNLKOIYYNSK-----AITSSEKAD-QFLTNTLLF 54
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 295 VNGVDMRKA-----SSHEYNNVEVPFETKLSRIGISSLEKSRNQLLNDILF 343

Qy 55 KGFTTGHWPYNDLLVDLG-----STAATSEYEGSSVDLYGAYGYOCAGGTPNKACMYG 109
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 344 -----NNALTDLNLKSLATSIERYKNSNITM-----KTKT----- 375

Qy 110 GVTLHDNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHLYHGKFGCL 169
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 376 -LNQDESTVNEPKT-----RRPLLIDREKFLNKEFLDEIAREIYEFTLSEFKD 425

Qy 170 YNSDS 174
      |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |  :  |
Db 426 LNSDT 430

RESULT 42
US-10-014-436-4
; Sequence 4, Application US/10014436
; Publication No. US20020182699A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, Yuji
; APPLICANT: IGARASHI, Kazuaki
; APPLICANT: OZAKI, Katsuya
; APPLICANT: ARA, Katsutoshi
; APPLICANT: KAWAI, Shuji
; APPLICANT: ITO, Susumu
; TITLE OF INVENTION: Gene for Enzyme Having Both Alkaline Pullulanase and Alkaline
; FILE OF INVENTION: Amylase Activities
; FILE REFERENCE: 2173-0122P
; CURRENT APPLICATION NUMBER: US/10/014,436
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: JP 111547/1995
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: PCT/JP96/01243
; PRIOR FILING DATE: 1996-05-10
; PRIOR APPLICATION NUMBER: US 08/952,084
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: US 09/514,302
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 893
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-10-014-436-4
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 8, 2003, 11:55:11 ; Search time 8.75359 Seconds
(without alignments)
2558.872 Million cell updates/sec

Title: US-09-900-766-2
Perfect score: 1218
Sequence: 1 SEKSEINEKDLRKSELOQ.....RDNTTISSTLSLSLYLTT 233

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1107	90.9	257	2	enterotoxin E prec
2	948	77.8	257	2	enterotoxin A prec
3	918	75.4	260	2	enterotoxin P [imp
4	611	50.2	258	2	enterotoxin D prec
5	420.5	34.5	258	2	enterotoxin Sen [i
6	372.5	30.6	260	2	enterotoxin Se0 [i
7	311	25.5	240	2	extracellular ente
8	296.5	24.3	242	2	extracellular ente
9	284.5	23.4	266	2	enterotoxin C3 - S
10	282	23.2	239	2	enterotoxin SEM [i
11	277.5	22.8	266	2	enterotoxin C-2 pr
12	269.5	22.1	251	1	exotoxin type A pr
13	268.5	22.0	266	1	enterotoxin B prec
14	263.5	21.6	266	1	enterotoxin C-1 pr
15	253.5	20.8	236	2	exotoxin type A pr
16	251.5	20.6	236	2	exotoxin type A pr
17	248.5	20.4	236	2	extracellular ente
18	241	19.8	258	2	streptococcal pyro
19	220	18.1	250	1	enterotoxin YENT2
20	190	15.6	136	2	hypothetical prote
21	178	14.6	157	2	exotoxin C precurs
22	155.5	12.8	235	2	exotoxin C precurs
23	115	9.4	62	2	hypothetical prote
24	100	8.2	234	1	toxic shock syndro
25	98	8.0	133	2	enterotoxin Yent1
26	98	8.0	227	2	exotoxin 14 [impor
27	98	8.0	234	2	toxic shock syndro
28	96	7.9	552	2	chitinase chi-A or
29	94	7.7	231	2	exotoxin 11 [impor

ALIGNMENTS

RESULT 1

A28179
enterotoxin E precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 15-Oct-1999
C:Accession: A28179
R:Couch, J.L.; Soltis, M.T.; Betley, M.J.
J. Bacteriol. 170, 2954-2960, 1988
A:Title: Cloning and nucleotide sequence of the type E staphylococcal enterotoxin gen
A:Reference number: A28179; MUID:88257005; PMID:3384800
A:Accession: A28179
A:Molecule type: DNA
A:Residues: 1-257 <COU>
A:Cross-references: GB:M21319; NID:gl53001; PIDN:AAA26617.1; PID:gl53002
C:Superfamily: enterotoxin B

Query Match 90.9%; Score 1107; DB 2; Length 257;
Best Local Similarity 89.7%; Pred. No. 9.8e-88;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELOQTALGNLKQIYYNSKAITSEKSADQFLNTLFLKGFFTG 60
DB 25 SEKSEINEKDLRKSELOQTALGNLKQIYYNSKAITSEKSADQFLNTLFLKGFFTG 84
QY 61 HPWYNDLLVDLGSTATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
DB 85 HPWYNDLLVDLGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTVTPIDKVKTSKEVTVOELDQARHYLHGKFGLYNSDSFGGKVQ 180
DB 145 EEKVPINLWIDGKQTVTPIDKVKTSKEVTVOELDQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTSVSYDLFDAQOQYPTDLLRIYRDNTTISSTLSLSLYLTT 233
DB 205 RGLIVFHSSEGSTSVSYDLFDAQOQYPTDLLRIYRDNTTISSTLSLSLYLTT 257

RESULT 2

A28664
enterotoxin A precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 15-Oct-1999
C:Accession: A28664; A29566
R:Betley, M.J.; Mekalanos, J.J.
J. Bacteriol. 170, 34-41, 1988
A:Title: Nucleotide sequence of the type A staphylococcal enterotoxin gene.
A:Reference number: A28664; MUID:88086892; PMID:3335483
A:Accession: A28664
A:Molecule type: DNA
A:Residues: 1-257 <BET>
A:Cross-references: GB:M18970; NID:gl53120; PIDN:AAA26681.1; PID:gl53121
A:Experimental source: strain FRI337

Db 92 SSKVFNSSDLANOFKGNKNDIYGLYFGNKCGLTEKTSCLYGVTTIHDGNQDDEKV 151
QY 125 VPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFLYNSDFSFGKVQORGLI 184
Db 152 IGWVFKDGVQEGFV--IKTKAKVTVOELDTKVRKLENIYKIYNKDT--GNIOKGCI 207
QY 185 VFHS--SEGTSVSYDLFDAQGOYPTDLLRIYRDNTTISSTLSLSILYLY 231
Db 208 FFHSHHQDSFYDLYNKGSGVGAFFQFYSDNRTVSSSNHYHIDVFLY 256

RESULT 6
E89969
enterotoxin se0 [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: E89969
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: E89969
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701623; PIDN:BAB42916.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: se0
C:Superfamily: enterotoxin B

Query Match 30.6%; Score 372.5; DB 2; Length 260;
Best Local Similarity 38.8%; Pred. No. 1.5e-24;
Matches 93; Conservative 32; Mismatches 92; Indels 23; Gaps 7;

QY 8 NEKD-----LRKSKELQGTALGNLKOIYYNS-KAITSSEKSAQDLTNTLLFKGFFTG 61
Db 29 NEEDPKIESLCKSSVDPTALHNDIYNNRFTTVKSVSTTEKFLDPLFKSI---- 84

QY 62 PW-----YNDLLVDLGSTAATSEYEGSSVDLYGAYGYOCAGTTPNKTACMGVTLH 114
Db 85 NWLDGTSAEFKDLKVEFSSAISKEFLGKTVDIYGVYKHAHGEHQVDCTACTYGGVTPH 144

QY 115 DNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFLYNS 174
Db 145 ENKUSEPKNIGVAVYKDNVNTFI--VTTDKKVTQAQELDIKVRTKLNNAYKLY--DR 200

QY 175 FGKVGQGLIVFHSSEGSTVS--YDLFAQGOYPTDLLRIYRDNTTISSTLSLSILYLT 232
Db 201 MTSQVQKGYKFHSHSEKHSFYDLFIYKGNLPOYLQIYNDKNTIIDSDDHYDVLFT 260

RESULT 7
G89991
extracellular enterotoxin L [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89991
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G89991
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-240 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701803; PIDN:BAB43096.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:

A:Gene: sel
C:Superfamily: enterotoxin B

Query Match 25.5%; Score 311; DB 2; Length 240;
Best Local Similarity 33.3%; Pred. No. 2.6e-19;
Matches 74; Conservative 36; Mismatches 78; Indels 34; Gaps 6;

QY 24 GNLKQIY---YNSKAITSSSEKSAQDLTNTLLFKGFFTGHPWYND--LLVDLGSTAATS 78
Db 31 GNLRFYTKYEYVNLKNVKNKNSPESHRL-----YSYRNDTLYAEFDNEYITS 79

QY 79 EYEGSSVDLYGAYGYOCAGTTPNKTACMGVTLHDNNRLTEKKVPINLWIDGKQTTV 138
Db 80 DLKGNVDVFGISYKY---GNSRT--IYGGVTRAKNKLDSPRIPIINLIINGKHQTV 133

QY 139 PIDKVKTSKEVTVQELDLQARHYLHGKFLY-----NSDSFGGKVQGLIVF 186
Db 134 TPKSVSTDKMVTQAQELDIKRLKYLQDEFNHYGHNDTGKKEYGTSSKFYSGFDKGSVVF 193

QY 187 HSEGSTSVSYDLFDAQGOYPTDLLRIYRDNTTISSTLSLSISL 228
Db 194 HMDGNSFSYDLFYTGYGLPESFLKIYKDNKTVDSTQPHLDV 235

RESULT 8
C89969
extracellular enterotoxin type I precursor [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89969
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89969
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701621; PIDN:BAB42914.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: sei
C:Superfamily: enterotoxin B

Query Match 24.3%; Score 296.5; DB 2; Length 242;
Best Local Similarity 32.4%; Pred. No. 4.7e-18;
Matches 77; Conservative 39; Mismatches 85; Indels 37; Gaps 10;

QY 8 NEKDLRKSELOG-TALGNLKOIY----YNSKAITSSSEKSAQDLTNTLLFKGFFTGHP 62
Db 17 NIKDL---TYAGGDIGVGNLRFYTKHDYIDLKGVTDKNLP----IANQLEFS---TG-- 64

QY 63 WYNDLLVDLGSTAATSEYEGSSVDLYGAYGYOCAGTTPNKTACMGVTVTLHDNNRLTEE 122
Db 65 -TNDLISNSNWDSEKSKFKKKDLIFGIDYNGPC-----KSKYMYGGATL-SGOYLSA 116

QY 123 KKVPIINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFLYNSDFS-GR--- 178
Db 117 RKIPINLWVNGKHKTISTDKIATNKKLTQAQELDIKRLKYLQDEFNHYGHNDTGKKEYG 176

QY 179 -----VQRGLIVFHSSEGSTVSYDLFDAQGOYPTDLLRIYRDNTTISSTLSISL 228
Db 177 YKSIFYSGFNNGKVLFLHNLNNEKFSYDLFYTGDLGDLPSFLKIYEDNKITIESEKFLDV 234

RESULT 9
S11885
enterotoxin C3 - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Jul-1999
C:Accession: S11885

R:Hovde, C.J.; Hackett, S.P.; Bohach, G.A.

Mol. Gen. Genet. 220, 329-333, 1990

A:Title: Nucleotide sequence of the staphylococcal enterotoxin C3 gene: sequence comparison

A:Reference number: S11885; MUID:90220508; PMID:2325627

A:Accession: S11885

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-266 <HGV>

A:Cross-references: GB:X51661; NID:g46570; PIDN:CAA35972.1; PID:g46571

C:Superfamily: enterotoxin B

Query Match 23.4%; Score 284.5; DB 2; Length 266;
Best Local Similarity 32.5%; Pred. No. 5.8e-17;
Matches 76; Conservative 45; Mismatches 94; Indels 19; Gaps 9;

Qy 11 DLKSKSELGATGALNKLQIYYNKAITSSE--KSAQDFLTNTLLFKGFTTGPWVNDLLV 69

Db 37 DLHKSSEFTGT-MGNMK--YLYDDHYVSATKVKSVDRFLAHDLIYNISDKLKNYDKVKT 93

Qy 70 DLGSAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGGVTLHDNNRLTEE 122

Db 94 ELLNEDLAKKYKDEVDVYGSNYVNCYFSSKDNVGVKVTGKTCMYGGITKHEGHPDNG 153

Qy 123 --KKVPINLWDIGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLSNDSFGGKQV 180

Db 154 NLONVLRVY-ENKRNITISFE-VQTDKKSVAQELDIKARNFLINKKNLYEFNS--SPYE 209

Qy 181 RGLIVFHSSEGSTVSYDLFDAQGYPD--TLRIYRDNTTISSTLSLSLYLT 232

Db 210 TGYIKFIENNGNTFYDMMMPAGDKFDOSKYLMMYNDKNTVDSKSVKIEVHLT 263

RESULT 10

D89969

enterotoxin SEM [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: D89969

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-239 <KUR>

A:Cross-references: GB:BA000018; PID:g13701622; PIDN:BA042915.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: sem

C:Superfamily: enterotoxin B

Query Match 23.2%; Score 282; DB 2; Length 239;

Best Local Similarity 29.7%; Pred. No. 8.2e-17;

Matches 68; Conservative 40; Mismatches 77; Indels 44; Gaps 6;

Qy 23 LGNLKQIYYNKAIT--SSEKSAQDFLTNTLLPKGFTTGPWVNDLLVGLGSAATSE- 79

Db 24 VGVNLNRNYGSPYEDHOSHPENNHLHQLVF-----SMDNSTVTAEF 68

Qy 80 -----YEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEKKVPINLWI 131

Db 69 KNVDDVKFKHADVYGLSVGYCL-----KNKYIYGGVTL-AGDYLEKSRPPINLWV 122

Qy 132 DGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGY-----NSDSFGGKV 179

Db 123 NGEHOTISTDKVSTNKKLVTAQETDKLRRYLOEYNIYFNDTKGRNYGNKSKFSSGF 182

Qy 180 QRLIVFHSSEGSTVSYDLFDAQGYPD--TLRIYRDNTTISSTLSLSL 228

Db 183 NAGKILFLHNDGSSFSYDLFTGTGTGQAEFLKIYNDKNTVETKFDLV 231

RESULT 11

A60114

enterotoxin C-2 precursor - Staphylococcus aureus

N:Alternate names: enterotoxin C-3 precursor

C:Species: Staphylococcus aureus

C:Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 16-Jul-1999

C:Accession: A60114; B60114; A33866

R:Bohach, G.A.; Schlievert, P.M.

Infect. Immun. 57, 2249-2252, 1989

A:Title: Conservation of the biologically active portions of staphylococcal enterotoxin

A:Reference number: A60114; MUID:89277549; PMID:2543637

A:Accession: A60114

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-266 <BOH>

A:Accession: B60114

A:Molecule type: protein

A:Residues: 28-66 <BOH2>

R:Couch, J.L.; Betley, M.J.

J. Bacteriol. 171, 4507-4510, 1989

A:Title: Nucleotide sequence of the type C-3 staphylococcal enterotoxin gene suggests

A:Reference number: A33866; MUID:89327174; PMID:2473979

A:Accession: A33866

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-266 <COU>

A:Cross-references: GB:M28364; NID:g153003; PIDN:AAA26624.1; PID:g153004

C:Genetics:

A:Gene: entC2

C:Superfamily: enterotoxin B

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-266/Product: enterotoxin C-2 #status predicted <MAT>

Query Match 22.8%; Score 277.5; DB 2; Length 266;

Best Local Similarity 31.1%; Pred. No. 2.3e-16;

Matches 76; Conservative 47; Mismatches 102; Indels 19; Gaps 9;

Qy 1 SEKSEINEKDLRKKSELOGTALGNLQIYYNKAITSSE--KSAQDFLTNTLLPKGFTT 59

Db 27 AESQPDPTPELHKSSEFTGT-MGNMK--YLYDDHYVSATKVKSVDRFLAHDLIYNISDK 83

Qy 60 GHPWYNOLLVDLGSTAAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGVT 112

Db 84 KLKNYDKVKTTELLNEDLAKKYKDEVDVYGSNYVNCYFSSKDNVGVKVTGKTCMYGGIT 143

Qy 113 LHDNNRLTEE--KKVPINLWDIGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGY 170

Db 144 KHEGHPDNGNLQNVLRVY-ENKRNITISFE-VQTDKKSVAQELDIKARNFLINKKNLY 201

Qy 171 NSDSFGGKVQRLIVFHSSEGSTVSYDLFDAQGYPD--TLRIYRDNTTISSTLSLSL 228

Db 202 EFNS--SPYETGYIKFIENNGNTFYDMMMPAGDKFDOSKYLMMYNDKNTVDSKSVKIEV 259

Qy 229 YLYT 232

Db 260 HLT 263

RESULT 12

S29659

exotoxin type A precursor (allele 1) - Streptococcus pyogenes phase T12

N:Alternate names: erythrogenic toxin; scarlet fever toxin

C:Species: Streptococcus pyogenes phase T12

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S29659; S18782; S18784; S18785; S18791; S18796; S18800

R:Weeks, C.R.; Ferretti, J.J.

Infect. Immun. 52, 144-150, 1986

A:Title: Nucleotide sequence of the type A streptococcal exotoxin (erythrogenic toxin

A:Reference number: S29659; MUID:8616804; PMID:3514452

A:Accession: S29659

A:Molecule type: DNA

C:Species: Streptococcus pyogenes
A:Variety: strain MGAS262 isolate California
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S18789
R:Nelson, K.; Schlievert, P.M.; Selander, R.K.; Musser, J.M.
J. Exp. Med. 174, 1271-1274, 1991
A:Title: Characterization and clonal distribution of four alleles of the speA gene encoded by the enterotoxin A of Streptococcus pyogenes
C:Reference number: S18782; PMID:92044323; PMID:1940804
A:Accession: S18789
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61573; NID:g47303; PIDN:CAA43771.1; PID:g47304
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
C:Genetics:
A:Gene: speA
C:Superfamily: enterotoxin B
C:Keywords: exotoxin
F:1-22/Domain: signal sequence (fragment) #status predicted <SIG>
F:23-236/Product: exotoxin A (fragment) #status predicted <MAT>

Query Match 20.6%; Score 251.5; DB 2; Length 236;
Best Local Similarity 31.7%; Pred No. 3.4e-14;
Matches 72; Conservative 42; Mismatches 102; Indels 11; Gaps 7;

QY 4 SEINEKDLRKXSELOGTAL-GNLKOIY--YNSKAITSEKSADQFLTNTLLFKGFFTG 60
Db 17 SQEVFAQQDPNPQLSRSLVKNLQNIYFLYEGDPVHENVKSVQDQLSHDLIYN--VS 73

QY 61 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYQC-AGTPNKTACMGVTLHDNNRL 119
Db 74 GLNYDKLKTALNEMSTFKNKNDIYVEVYHCLCRNAKRACIYGGVTNHEGNNHL 133

QY 120 TEEKKVPINLWDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKV 179
Db 134 EIPKNTLVKVSIDGQ-SLSFD-IETSKMVTQAQLDYKVRKHLTDNKKQLYNGP--SKY 189

QY 180 QRGLIVFHSSEGSTVSYDLFDAGQGYPDTLRLIYRDNTTISSTLSI 226
Db 190 ETGYIKFISKDKETWFDFPEPEFQNVKLYMKDKNETLDSSTSOI 236

RESULT 17
S18786
N:Alternate names: scarlet fever toxin
C:Species: Streptococcus pyogenes phage
A:Variety: strain MGAS250 isolate California; strain MGAS251 isolate California; strain MGAS252 isolate California
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S18786; S18787; S18788; S18790; S18792; S18795; S18799
R:Nelson, K.; Schlievert, P.M.; Selander, R.K.; Musser, J.M.
J. Exp. Med. 174, 1271-1274, 1991
A:Title: Characterization and clonal distribution of four alleles of the speA gene encoded by the enterotoxin A of Streptococcus pyogenes
C:Reference number: S18782; PMID:92044323; PMID:1940804
A:Accession: S18786
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61561; NID:g47297; PIDN:CAA43759.1; PID:g47298
A:Experimental source: strain MGAS250 isolate California unassigned phage
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18787
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61562; NID:g47299; PIDN:CAA43760.1; PID:g47300
A:Experimental source: strain MGAS251 isolate California unassigned phage
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18788
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>

A:Cross-references: EMBL:X61563; NID:g47301; PIDN:CAA43761.1; PID:g47302
A:Experimental source: strain MGAS256 isolate California unassigned phage
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18790
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61564; NID:g47305; PIDN:CAA43762.1; PID:g47306
A:Experimental source: strain MGAS285 isolate Colorado unassigned phage
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18792
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61565; NID:g47311; PIDN:CAA43763.1; PID:g47312
A:Experimental source: strain MGAS480 isolate Yugoslavia unassigned phage
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18795
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61566; NID:g47317; PIDN:CAA43764.1; PID:g47318
A:Experimental source: strain MGAS492 isolate United Kingdom unassigned phage
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18799
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61567; NID:g47325; PIDN:CAA43765.1; PID:g47326
A:Experimental source: strain MGAS496 isolate Germany unassigned phage
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
C:Genetics:
A:Gene: speA2
C:Superfamily: enterotoxin B
C:Keywords: exotoxin
F:1-22/Domain: signal sequence (fragment) #status predicted <SIG>
F:23-236/Product: exotoxin type A (fragment) #status predicted <MAT>

Query Match 20.4%; Score 248.5; DB 2; Length 236;
Best Local Similarity 31.6%; Pred No. 6.1e-14;
Matches 73; Conservative 44; Mismatches 95; Indels 19; Gaps 10;

QY 4 SEINEKDLRKXSELOGTAL-GNLKOIY-YNSKAIT-SEKSADQFLTNTLLFKGFFTG 60
Db 17 SQEVFAQQDPNPQLSRSLVKNLQNIYFLYEGDPVHENVKSVQDQLSHDLIYN--VS 73

QY 61 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYQC-AGTPNKTACMGVTLHDNNRL 119
Db 74 GPNYDKLKTALNEMSTFKNKNDIYVEVYHCLCRNAERACIYGGVTNHEGNNHL 133

QY 120 TEEKKVPINLWDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKV 179
Db 134 EIPKNTLVKVSIDGQ-SLSFD-IETSKMVTQAQLDYKVRKHLTDNKKQLYNGP--SKY 189

QY 180 QRGLIVFHSSEGSTVSYDLFD---AQGYPDTLRLIYRDNTTISSTLSI 226
Db 190 ETGYIKFIPKNESEWFDFPEPEFTQSKY---LMIYKDNELTDSNTSOI 236

RESULT 18
G89968
extracellular enterotoxin type G precursor [imported] - Staphylococcus aureus (strain
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89968
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: G89968
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-258 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701617; PIDN:BAB42910.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: seg
C:Superfamily: enterotoxin B

Query Match 19.8%; Score 241; DB 2; Length 258;
Best Local Similarity 28.8%; Pred. No. 3e-13; Mismatches 88; Indels 42; Gaps 12;
Matches 72; Conservative 48

QY 3 KSEINEKDLRKSEKQGTALGNLKOIYY---YNSKAITSSSEKSAQDQFLNTLLFKGFFT 59
Db 30 KLDLKNKVSQKNN--KGT-MGNVNLVYTPPVGGRGVNSR-----QFLSHDLIFP---I 79
QY 60 GHPWYNDLLVGLGSTAATSEYEGSSVDLYGAYYGQCA-----GGTPFNKTACM 107
Db 80 EYKSYNEVKTLENTLANNYKDKVDIFGVFYFTCIIPKSPDINQNFSG---CCM 134
QY 108 YGGVTLH--DNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHG 165
Db 135 YGGLTNSSENER--DKLTVQVTTIDNRQSLG--FTITNKNMVTIQEUDYKARHWTK 189
QY 166 KFGLYNSDFGGKVQRLVHFHSSEGSTSYDLFDAQGOYPDV---LLRIYRDNTTISST 222
Db 190 EKKLYEFD--GSAFESGYIKFTEKNTSFWDLPFKKELVPFPVYKFLNYGDNKVVDK 247
QY 223 SLISILYLT 232
Db 248 SIKMEVFLNT 257

RESULT 19
A26152
streptococcal pyrogenic exotoxin type A precursor - Streptococcus sp.
N:Alternate names: scarlet fever toxin; SPE type A (speA)
C:Species: Streptococcus sp.
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A26152
R:Johnson, L.P.; L'Italien, J.J.; Schlievert, P.M.
Mol. Gen. Genet. 203, 354-356, 1986
A:Title: Streptococcal pyrogenic exotoxin type A (scarlet fever toxin) is related to sta
A:Reference number: A26152; MUID:86284313; PMID:3526093
A:Accession: A26152
A:Molecule type: DNA
A:Residues: 1-250 <JOH>
C:Superfamily: enterotoxin B
C:Keywords: exotoxin

Query Match 18.1%; Score 220; DB 1; Length 250;
Best Local Similarity 28.9%; Pred. No. 1.9e-11;
Matches 66; Conservative 42; Mismatches 98; Indels 22; Gaps 9;

QY 14 KXSELOGTALGNLKOIY-YNSKAITSSE--KSADQFLNTLLFKGFFTGHWPYNDLLVD 70
Db 35 KPSQLQRSNLVTKYIIFRMVTLVTHENVKSVQDQLLSDLIYN---VSGPNYDKLKTE 91
QY 71 LGSTAATSEYEGSSVDLYGAYYGQC-AGTPNKTACMYGGVTLHDNNRLTEKKVPINL 129
Db 92 LKNQEMATLFDKNVDIYGVEYHLCYLCENAPRSACLYGGVTNHEGNHLEPKKIVKRV 151
QY 130 WIDGKQT-TVPIDKVKTSKKEVTQELDLQARHYLHGKFLYNSDSFGGKVQRLVHFS 188
Db 152 SIDGIQSLFSDIEQKNG----NCSRISTYVRKYLTDNKLTYNGP--SKYETGYIKFIP 205
QY 189 SEGSTSVYDLFD----AQOQYPDTLRIYRDNTTISSTSLISLYLT 232
Db 206 KNKSEWFDFPEPTQSKY----LMYKDNTELDTSNTSQIEVLYLT 249

RESULT 20
A89969

enterotoxin YENT2 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: A89969
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: A89969
A:Molecule type: DNA
A:Residues: 1-136 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701619; PIDN:BAB42912.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: yent2

Query Match 15.6%; Score 190; DB 2; Length 136;
Best Local Similarity 34.1%; Pred. No. 3.2e-09;
Matches 46; Conservative 25; Mismatches 54; Indels 10; Gaps 5;
QY 103 KTACMYGGVTLHDNNRLTEKKV--PINLWI---DGKQTTVPIDKVKTSKKEVTVOELDL 157
Db 2 KKTCTMYGGVTEHGDGQIDKNNSTDNSHNLIKVYENRNSLSFD-IPTNKKNITAEIDY 60
QY 158 QARHYLHGKFLYNSDSFGGKVQRLVHFHSSEGSTSVYDLFDAQGO--YPTDLRIYRD 215
Db 61 KVRNYLLKHKHLYEFNS--SPYETGYIKFIEGSHSFWDLPESGKKFYPTKYLIIYND 118
QY 216 NTTISSTSLISLYL 230
Db 119 NKTVESKSNVEVHL 133

RESULT 21
A89942
hypothetical protein SA1430 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: A89942
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: A89942
A:Molecule type: DNA
A:Residues: 1-157 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701400; PIDN:BAB42694.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1430

Query Match 14.6%; Score 178; DB 2; Length 157;
Best Local Similarity 31.4%; Pred. No. 4.2e-08;
Matches 37; Conservative 21; Mismatches 58; Indels 2; Gaps 1;

QY 1 SKSEIEKDKRLKSKLQGTALGNLKOIYYVNSKAITSSSEKSAQDQFLNTLLFKGFFT 60
Db 42 TNSAIEYSDLHHKSKFDSKFLSNKSNFINTQ--LENKNTNDRLLKHLDFHDMFVN 99
QY 61 HPWYNDLLVGLGSTAATSEYEGSSVDLYGAYYGQCAGTPNKTACMYGGVTLHDNNR 118
Db 100 VASKKDKFVEPENEALSKKFINKNDIIVAGSYSECHGATNKTCCSYGGVTLSDNNK 157

RESULT 22
A30509
exotoxin C precursor - Streptococcus pyogenes

J. Biol. Chem. 261, 15783-15786, 1986

A::Title: The nucleotide and partial amino acid sequence of toxic shock syndrome toxin A:
A:Reference number: A24606; MUID:87057222;. PMID:3782090
A:Accession: A24606
A:Molecule type: DNA
A:Residues: 1-234 <BL>
A:Cross-references: GB:J02615; NID:g153122; PIDN:AAA26682.1; PID:g153123
A:Accession: B24606
A:Molecule type: protein
A:Residues: 41-106;199-224 <BLO2>
A>Note: the authors translated the codon ACT for residue 20 as Ile and GGG for residu
C:Superfamily: toxic shock syndrome toxin
C:Keywords: toxic shock syndrome: toxin
F:I=40/Domain: signal sequence #status predicted <SIG>
F:I=234/Product: toxin shock syndrome toxin-1 #status experimental <MAT>

Query Match 8.2%: Score 100; DB 1; Length 234;
Best Local Similarity 24.5%; Pred.No.0.36;
Matches 48; Conservative 33; Mismatches 61; Indels 54; Gaps 12;

Qy 25 NLKQT---YYNSKAITSSE-----KSADQLTNTLLFKGFTGHWPYNL-- 68
 | : | | | | |
Db 45 NIKDLDWYSGSDFTNSEVLDSLGRMKIKNDGSI-SLIIFPS----PYSPAFTK 98

Qy 59 ---VDLGSTAATSEYGSSVLYGAYGYQCAGGTAKMYCGGVTLHNNRUTEKKV 125
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 99 GEKVLD-NTRKTCKSQHTS---EGTYIHFIQSIVT-----NTEKLPTIEL 140

Qy 126 PNLWDIGDKOTVPIDKVKTSKEVTNVQEILDQRHYLHGKFGLY-NSDSFGG--KVORG 182
 | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 141 PLUKVVGKDDSLPKYGP-KFDKRQLAISTDLEIRHQLTQHGLYRSSDRTGGYWKIT-- 197

Qy 183 LIVFHSSSEGSTVSIDL 198
 ::||| ||
Db 198 -----MNDGSTYQSDL 208

RESULT 25
B89969
enterotoxin Yentl [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89969
E:Kuroda, M.; Ohta, T.; Uchiyama, I.: Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ut, Y.; Kobayashi, N.; Sawano, R.; Inoue, R.; Kaito, C.; Sekimizu,
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A>Title: Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*.
A:Reference number: AB9758; MUID:21311952; PMID:11418146
A:Accession: B89969
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-133 <KUR>
A:Cross-references: GB:BA000018; PID:g13701620; PIDN:BABA2913.1; GSPDB:GN00149
A:Experimental source: strain N315

```

C:Genetics:
A:Gene: yent1

      8.0%:  Score 98;  DB 2;  Length 133;
Best Local Similarity 26.2%;  Pred. No. 0.25;
Matches 34;  Conservative 27;  Mismatches 41;  Indels 28;  Gaps 8;

Qy  3  KSEINEKDLRKKSELOQTALGNLKOYYYNKAITSS- KSAQDFLTNTLLFKGFFTG 61
      | | | | | | | | | | | | | | | | | | |
Db  25  KPEQLN-----KASEFTG-LMDNMR--YLYDDKHVSETNIKSOEKFQHLDLFK--INGS 74

Qy  62  PWNLLVLDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTE 121
      | | | | | | | | | | | | | | | | | | |
Db  75  ---KILKTEFNKNSLSDKYKNKNVDLFGTYNNQC-----YFSL---DNMELND 117

Qy  122  EKKVPINLWI 131
      : | | |
Db  118  GLIEKNVYV 127

```

```

RESULT 26
G89807
exotoxin 14 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89807
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G89807
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-227 <KUR>
A:Cross-references: GB:BA000018; PID:g13700320; PIDN:BAB41618.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: set14

Query Match      8.0%; Score 98; DB 2; Length 227;
Best Local Similarity 26.4%; Pred. No. 0.51;
Matches 53; Conservative 28; Mismatches 78; Indels 42; Gaps 9;

QY 8 NEKDRLKSELQGTALGNLKQIY-YNNSKAITSSSEKADQFTNTLLFKGFTGHPWYND 66
DB 33 NOKSVNKHDK-----EALRYTGTMTKMKNISALKHGKNNLRFK--FRGIKIQVL 81
QY 67 LLVDLGSTAATSEYEGSSV-----DLYGAYGYQCAGGT--PNKTACMGYGVTLHDNN 117
DB 82 LPGNDKSKFQORSYEGLDVFFVQERKHDIYF---TVGVTQNNKTS-----GVVSAPIL 134
QY 118 RUTEKPKVPINLWIDKQTTVPIDRKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGG 177
DB 135 NISKEK-----GEDAFVKGYPYIKKEKITLKELDYKURKHLEIKYGLYKTSKDG 185
QY 178 KVQRLIVFHSSEGSTVSYDL 198
DB 186 RVKISL-----KDGsfYNLDL 201

RESULT 27
B89992
toxic shock syndrome toxin-1 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89992
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89992
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <KUR>
A:Cross-references: GB:BA000018; PID:g13701806; PIDN:BAB43099.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: tst
C:Superfamily: toxic shock syndrome toxin

Query Match      8.0%; Score 98; DB 2; Length 234;
Best Local Similarity 24.6%; Pred. No. 0.54;
Matches 49; Conservative 33; Mismatches 57; Indels 60; Gaps 13;

QY 25 NLKQI---YYNNSKAITSE-----KSADQFTNTLLFKGFTGHPWYNDLL-- 68
DB 45 NIKDLDWYSSGSDFTNSELVNSLGSRMKNTDGS1-SLIIFPS-----PYSPAFTK 98
```

```

QY 69 ---VDLGSTAATSEYEGSSVDLYGAYGYQCAGGTGNKTACMGYGVTLHDNNRLTEKKV 125
DB 99 GEKVDL-NTRTKKKSQHTS---EGYIHFQISGVT-----NTEKLPPIEL 140
QY 126 PINLWIDKQTTV---PIDKVKTSKKEVTVOELDLQARHYLHGKFGLY-NSDSFGG--KV 179
DB 141 PLKVKRVHGKDSPLKYWP-----NFDKKQLAISTLDPEIRHQLTQIHGLYRSSDKTGYYWKI 196
QY 180 QRGLIVFHSSEGSTVSYDL 198
DB 197 T-----MNDGSTYQSGL 208

RESULT 28
T41863
chitinase chi-A orf126 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BMSNPV
A:Variety: isolate T3
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T41863
R:Gomi, S.; Wajima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: Z22020; MUID:99281911; PMID:10355780
A:Accession: T41863
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-552 <KAM>
A:Cross-references: EMBL:L31180; NID:g3745835; PIDN:AAC63792.1; PID:g3745945
A:Experimental source: isolate T3
C:Superfamily: Serratia marcescens chitinase

Query Match      7.9%; Score 96; DB 2; Length 552;
Best Local Similarity 24.9%; Pred. No. 2.5;
Matches 46; Conservative 23; Mismatches 68; Indels 48; Gaps 11;

QY 32 YNSKAITS---SEKSADQFTNTLLFKGFTGHPWYNLLVDLGSTAATSEYEGSSVDLY 88
DB 359 YDKIAVNVYAEAKSLDKIFLWYDFKG-----AWSN---TDLGYOTTVYVSWNSEELY 410
QY 89 GAYGYQC---AGGTPNKT---ACMYG-----GVTLDN-NRLTEKKVPIN-LWIDGKQT 136
DB 411 TTHYAVDALLEGQVDPNKTIIVGVAMVGRGWTGVTNTNGNYFSGTGNGPVGSGTWEDG--- 467
QY 137 TVPIDKVKTSKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVGORGLIVFHSSEGSTVSY 196
DB 468 -----VVDYRQIQKDLNNY-----VYTFDS-----AAQASYVFDKSKGDLISF 505
QY 197 DLFDA 201
DB 506 DSVDS 510

RESULT 29
D89807
exotoxin 11 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D89807
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D89807
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <KUR>
A:Cross-references: GB:BA000018; PID:g13700317; PIDN:BAB41615.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
```


Query Match 7.5%; Score 91; DB 2; Length 1634;
Best Local Similarity 21.6%; Pred. No. 27;
Matches 50; Conservative 36; Mismatches 79; Indels 66; Gaps 10;

QY 12 LRKKSELOCTALGNLKOIYYNSKAITSEKSADQFLTNTLFLKGFPTGHPWYNLLVDL 71
Db 614 VRRKGTTRAITLGCCKADLYLKIEIL---KNKEKYLDPNAIL-RGFFEGDGVNVRRAV 669

QY 72 GSTAATSEYE-----GSSVDLYGAYGYQCAGGTPNKTCACMYGGVTLHDNNRLTEKKVP 126
Db 670 VVNOGTNNYDKITFASLDRILGKYSF-----YTSYEERGGKLRVY 713

QY 127 INLWIDGKQTPVIDKVKTS-----KKEVTVOELDLQARHYLHGKFGFLYNSD----- 173
Db 714 IEIFSKG-----DLIKFSLISFISRRKNLLNEIIRQTKLYIKGIDFGFYLLDDVCVS 766

QY 174 --SFGKV-----QRGLVIFHSSEGS-TVSYDLDFDAQOQYPTL 209
Db 767 LESYKGEVVDLTLEGRPVYFANGILTHNSLYPSIIISYNI-----SPDTL 811

RESULT 33
C53312
thymidine phosphorylase (EC 2.4.2.4) - Mycoplasma pirum (strain BER)
C:Species: Mycoplasma pirum
C:Date: 23-Oct-1998 #sequence_revision 23-Oct-1998 #text_change 07-Dec-1999
C:Accession: C53312
R:Tham, T.N.; Ferris, S.; Kovacic, R.; Montagnier, L.; Blanchard, A.
J. Bacteriol. 175, 5281-5285, 1993
A:Title: Identification of Mycoplasma pirum genes involved in the salvage pathways for n
A:Reference number: A53312; MUID:93352438; PMID:8349569
A:Accession: C53312
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <THA>
A:Cross-references: GB:L13289; NID:g401781; PIDN:AAA25432.1; PID:g401784
C:Genetics:
C:Superfamily: thymidine phosphorylase
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 7.2%; Score 88; DB 1; Length 419;
Best Local Similarity 27.0%; Pred. No. 8.3;
Matches 38; Conservative 22; Mismatches 57; Indels 24; Gaps 7;

QY 47 FLTNTLLFKG-FFTGHWPYNLLVDLGGSTAATSE-----YEGSSVDLYGAYYG 93
Db 57 FLTKAMIDSGKIYKFPPEYKKILIDKHSTGGIGDKVSIARPLVDFDLGVAKLSGRGIG 116

QY 94 YOCAGGTPNKTCACMYGGVTLHDNNRLTEKKV--PINLWIDCK-QTTVPIDKVKTSKREV 150
Db 117 F--TGGTIDKLE----SINVTVDIDLKNSKKILNIANFIVGQTNDIVPADKLLYALRDV 170

QY 151 --TVQELDLQARHYLHGKFG 169
Db 171 TGTVDLSPLIARASILSKFAL 191

RESULT 34
A53310
pheromone CAD1 binding protein precursor - Enterococcus faecalis plasmid pAD1
N:Alternate names: TraC
C:Species: Enterococcus faecalis
C:Date: 08-Sep-1995 #sequence_revision 08-Sep-1995 #text_change 20-Aug-1999
C:Accession: A53310
R:Tanimoto, K.; An, F.Y.; Clewell, D.B.
J. Bacteriol. 175, 5260-5264, 1993
A:Title: Characterization of the traC determinant of the Enterococcus faecalis hemolysin
A:Reference number: A53310; MUID:93352433; PMID:8349566
A:Accession: A53310
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-543 <TAN>
A:Cross-references: GB:L19532; NID:g388267; PIDN:AAA72024.1; PID:g388269
C:Genetics:
A:Gene: traC
A:Genome: plasmid
C:Superfamily: dipeptide transport protein

Query Match 7.2%; Score 88; DB 2; Length 543;
Best Local Similarity 21.3%; Pred. No. 12;
Matches 46; Conservative 31; Mismatches 61; Indels 78; Gaps 10;

QY 33 NSKAITSEKSADQFLTNTLFLKGFPTG-----PWYNLLVDLGGSTAATSEY- 80
Db 146 NAKETASQKSQE-----TLAVKS--NGNKTIEIELEKPTPVFTDLL-----AUTAYYP 192

QY 81 -EGSSVDLYGAYGYQCAGGTPNKTCACMYGGVTLHDNNRLTEKKVPINLWIDGKQTTVP 139
Db 193 VQOKAIKEYGKY-----GTSQKSIVTNGAFNLATSLGCVGTSK-----W-----TIS 235

QY 140 IDKVKTSKKEVTVOELDLQARHYLHGKFGFLYNSDSFGGKVGQRLIVFHSSEGSTVSYDLF 199
Db 236 KNKEYWQREVSMEKINFQVKEINTGINLYN----- 267

QY 200 DAQGOYPTD-----LLRIYRDNTTISSTSLSLYL 230
Db 268 --DGOLDPTPTVGEYAKQYKONKEFSTTLTMANTMYL 301

RESULT 35
A71076
hypothetical protein PH0873 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999
C:Accession: A71076
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: A71076
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-788 <KAW>
A:Cross-references: GB:AP000004; NID:g3236131; PIDN:BAA29967.1; PID:d1030910; PID:g32
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa
C:Genetics:
A:Gene: PH0873

Query Match 7.2%; Score 88; DB 2; Length 788;
Best Local Similarity 23.3%; Pred. No. 19;
Matches 48; Conservative 36; Mismatches 66; Indels 56; Gaps 10;

QY 5 BEINEKDLRKRSELOCTALGNLKOIYYNSKAITSEKSADQFLTNTLFLKGFPTGHPWY 64
Db 180 EELNEENFRKELEEG--IKTLLEVAYYLKVI---DKENEYEINNLKPK----- 225

QY 65 NDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCACMYGGVTLHDNNRLTEKK 124
Db 226 NELKNDLSYLEASGYKYSQSQISQFYVALKTKG-----FVILSGLTSGTKTKIAQEL- 278

QY 125 VPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQA-----RH-----YLHGK----- 166
Db 279 -----GELLDLP--QVMSASGNTVAKREIKSLQETINRHRGFAYVGHWPFGKISKIK 328

QY 167 ----FGLYNS---DSFGGKVGQRLIV 185
Db 329 PPFIFWVYDSDENDYKKVPYGIIV 354

RESULT 36
AE1449
hypothetical protein lin0132 [imported] - Listeria innocua (strain Clip11262)

C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE1449
R:Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma, O.; Glaser, P.; Smeets, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1449
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1946 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC95365.1; PID:g16412551; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0132

Query Match 7.2%; Score 87.5; DB 2; Length 1946;
Best Local Similarity 22.0%; Pred. No. 69;
Matches 56; Conservative 30; Mismatches 98; Indels 71; Gaps 12;
QY 4 SEEINEKDLRKSELOQTALGNLKOIYY-----YNSKAITSSEKSADQFLTNL 53
DB 1198 TNLNLRADYRVTKGTPDSEGNAGVKFKNIWTKYIAYRGGMISNSETADPF-----S 1252
QY 54 FK----GFTGHPWYNLLVDLGSAATSEYEGSSVDLYGAYGYQCAGT-----P 101
DB 1253 FRQEDGSYVAB--UNGRYVDFSSGLVSTEAAGSKIDYSAQEGVSQVSAITTKALVEP 1310
QY 102 NKTACMYGVTLLHNNRLTEEEKVPIINLWIDGKQTPIDKVKTSKKEVTQ---ELDIQ 158
DB 1311 NITA---EVSIVDRR-----VKLDVVGEDKNTYTKKDNENELF 1347
QY 159 ARHYLHGFGLYNSDFSGKQVORGLIVPHSS-----EGSTVSYDLFDAQGYPTL 209
DB 1348 VNAYREDGETLVKSVRIDGLPTDFIVTDLSPDKYVIRKVEG---KYDLLDGEQ--PQN- 1401
QY 210 LRIYRDNNTTISSTSL 224
DB 1402 -KVVYSETITTEKSL 1415

RESULT 37
T08081
probable myrosinase-binding protein - rape
A:Alternate names: jasmonate inducible protein
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T08081
R:Geshi, N.; Brandt, A.
Planta 204, 295-304, 1998
A:Title: Two jasmonate-inducible myrosinase-binding proteins from Brassica napus L. seed
A:Reference number: Z16340; MUID:98192006; PMID:9530873
A:Accession: T08081
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-914 <GES>
A:Cross-references: EMBL:Y11482; NID:el023100; PIDN:CAA72270.1; PID:e304149
A:Experimental source: cv. Global; isolate a4; seedlings
A>Note: jasmonate inducible

Query Match 7.1%; Score 87; DB 2; Length 914;
Best Local Similarity 20.1%; Pred. No. 28;
Matches 65; Conservative 35; Mismatches 89; Indels 134; Gaps 14;
QY 2 EKSEINE-----KDLRKSELOQTALGNLKOIYY-YNS 34
DB 347 EKGEFNDVGEFGVKITVGADQYSVTIKIEYKDGKVEIREHGTSGELQEFSDVDPN 406
QY 35 KAITSSEKSADQFLT--NTLLFKGFTGHPWYNLLVDLGSTAATSEYE----- 81

Db 407 DSITEVGGTYKHNYYDTTLTSLYFTTSKGFTSPLFGINSEKKGTTEFEFKDENGKGLIG 466
QY 82 -----GSSVDLYGAYGYQCAGG-----TPNKTACMYGGVTLHDNNRLTEE 122
DB 467 LHGRGNAIDAIGAFDFTGSGQGDGVDVSKDGPWTDVPGKKGPLGG-----DKGEPPDD 521
QY 123 -----KKVPI---NLWIDGKQTPIDKVKTSKKEV-----TVQ 153
DB 522 VGFEGVKKITVGADNLSI---TYIKIEYKDGKVEIREHGTARGKIKKEFSVDYNDST 577
QY 154 ELDLARH-----YLHGKFGLYN-----SDSFGKGVORGLIV 185
DB 578 EVGGTYKHNYYDTTLTSLYFTTSKGFTSPLFGIDSEKKGTTEFEFKDENGK- ---LIG 633
QY 186 FHSSEGSTVSVDLFDAGQGYPT 208
DB 634 FHGRGNAI-----DAIGAYFDT 651

RESULT 38
T08102
myrosinase-binding protein (clone indmbp2) - rape
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 21-Jul-2000
C:Accession: T08102
R:Taipalensuu, J.; Eriksson, S.J.; Rask, L.
Eur. J. Biochem. 250, 680-688, 1997
A:Title: The myrosinase binding protein from Brassica napus seeds possesses lectin ac
A:Reference number: Z16355; MUID:98121188; PMID:9461290
A:Accession: T08102
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-988 <TAI>
A:Cross-references: EMBL:Y09437; NID:gl711295; PIDN:CAA70587.1; PID:gl711296
A:Experimental source: cv. Hanna; leaves

Query Match 7.1%; Score 87; DB 2; Length 988;
Best Local Similarity 20.5%; Pred. No. 31;
Matches 66; Conservative 35; Mismatches 89; Indels 132; Gaps 15;
QY 2 EKSEINE-----KDLRKSELOQTALGNLKOIYY-YNS 34
DB 363 EKGEFNDVGEFGVKITVGADQYSVTIKIEYKDGKVEIREHGTSGELQEFSDVDPN 422
QY 35 KAITSSEKSADQFLT--NTLLFKGFTGHPWYNLLVDLGSTAATSEYE----- 81
DB 423 DSITEVGGTYKHNYYDTTLTSLYFTTSKGFTSPLFGINSEKKGTTEFEFKDENGKGLIG 482
QY 82 -----GSSVDLYGAYGYQCAGG-----TPNKTACM-----YGGVTLHDNNRLTEE- 122
DB 483 LHGRGNAIDAIGAFDFTGSGQGDGVDVTKDVLADVPKKGKPLGG-----DKGEPPDDV 538
QY 123 -----KKVPI---NLWIDGKQTPIDKVKTSKKEV-----TVQE 154
DB 539 GFEGVKKITVGADNLSI---TYIKIEYKDGKVEIREHGTARGKIKKEFSVDYNDST 594
QY 155 LDLARH-----YLHGKFGLYN-----SDSFGKGVORGLIV 186
DB 595 VGGTYKHNYYDTTLTSLYFTTSKGFTSPLFGIDSEKKGTTEFEFKDENGK- ---LIG 650
QY 187 HSSEGSTVSVDLFDAGQGYPT 208
DB 651 HGRGNAI-----DAIGAYFDT 667

RESULT 39
E71403
hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: E71403

[illegible]

A: Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AAC05156.1
C: Genetics:
A: Gene: qtfn

N:Alternate names: pore-forming protein; protein D2
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 22-Nov-1993 #sequence:revision 10-Nov-1995 #text_change 15-Jun-2001
 C:Accession: S23771; S61560; S25008; E83527
 R:Yoneyama, H.; Yoshihara, E.; Nakae, T.
 Antimicrob. Agents Chemother. 36: 1791-1793, 1992
 A:Title: Nucleotide sequence of the protein D2 gene of *Pseudomonas aeruginosa*.
 A:Reference number: S23771; MUID:93037310; PMID:1339257
 A:Accession: S23771
 A:Molecule type: DNA
 A:Residues: 1-443 <YON>
 A:Cross-references: EMBL:X63152; NID:g45370; PIDN:CAA44855.1; PID:g45371
 A:Accession: S61560
 A:Molecule type: protein
 A:Residues: 24-34 <YON>
 R:Huang, H.; Siennel, R.J.; Bellido, F.; Rawling, E.; Hancock, R.E.W.
 submitted to the EMBL Data Library, July 1992
 A:Description: Analysis of two gene regions involved in the expression of the impA gene
 A:Reference number: S23859
 A:Accession: S25008

A:Residues: 1-443 <HDA>
A:Cross-references: EMBL:Z14065; NID:945372; PIDN:CAA78448.1; PID:945373
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.: Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: E83527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <STO>
A:Cross-references: GB:AE004529; GB:AE004091; NID:99946851; PIDN:AAG04347.1; GSPDB:CN
A:Experimental source: strain PA01
A:Genetics:
A:Gene: oprD; PA0958
C:Superfamily: *Pseudomonas aeruginosa* outer membrane porin PA0958
C:Keywords: porin
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-443/Product: protein D2 #status experimental <MAT>

Query Match 6.9%; Score 84.5; DB 2; Length 443;
Best Local Similarity 23.2%; Pred. No. 18;
Matches 60; Conservative 30; Mismatches 98; Indels 7

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:52:22 ; Search time 4.63425 seconds
(without alignments)
2085.338 Million cell updates/sec

Title: US-09-900-766-2
Perfect score: 1218
Sequence: 1 SEKSEINEKDLRKKSELQ.....RDNTTISTSLSLYLTT 233

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1107.	90.9	257	1 ETXE_STAAU	P12993 staphylococ
2	948	77.8	257	1 ETXA_STAAW	P13163 staphylococ
3	611	50.2	258	1 ETRD_STAAU	P20723 staphylococ
4	284.5	23.4	266	1 ETC3_STAAU	P23313 staphylococ
5	277.5	22.8	266	1 ETC2_STAAU	P34071 staphylococ
6	269.5	22.1	251	1 SPEA_STRPY	P08095 streptococ
7	268.5	22.0	266	1 ETXB_STAAU	P01552 staphylococ
8	263.5	21.6	266	1 ETC1_STAAU	P01553 staphylococ
9	241	19.8	258	1 ETRG_STAAU	O85382 staphylococ
10	185	15.2	236	1 SPEH_STRPY	O9X5C8 streptococ
11	183.5	15.1	234	1 SPEG_STRPY	O9X5C7 streptococ
12	148.5	12.2	235	1 TSST_STAAU	P13380 streptococ
13	100	8.2	234	1 TSST_STAAU	P08886 staphylococ
14	95.5	7.8	150	1 R19E_THEVO	Q97CU4 thermoplas
15	93	7.6	979	1 P115_MYCHR	P41508 mycoplasma
16	91	7.5	1634	1 DPOL_METUA	Q58295 methanococ
17	88.5	7.3	150	1 R19E_THEAC	Q9LM21 thermoplas
18	88	7.2	419	1 TVPH_MYCPT	P47717 mycoplasma
19	86	7.1	1388	1 HRP3_SCHPO	O14139 schizosach
20	84.5	6.9	443	1 PORD_PSEAE	P32722 pseudomonas
21	84	6.9	590	1 DNK1_STRMU	O06942 streptococ
22	84	6.9	1050	1 ITA5_XENLA	Q06274 xenopus lae
23	83.5	6.9	282	1 PRTA_ASPNG	P24665 aspergillus
24	82.5	6.8	328	1 YB70_HAEIN	O57527 haemophilus
25	82.5	6.8	455	1 YWAD_BACSU	P25152 bacillus su
26	82.5	6.8	600	1 NUCD_ECOLI	P33599 escherichia
27	82	6.7	607	1 DNK1_STRPY	P95831 streptococ
28	82	6.7	843	1 PULA_THEMA	O33840 thermotoga
29	82	6.7	3718	1 LMA5_MOUSE	Q61001 mus musculu
30	81	6.7	609	1 DNK1_STRAG	P95693 streptococ
31	81	6.7	611	1 WDR1_CAEEL	Q11176 caenorhabd
32	80	6.6	345	1 PURA_METUA	O57981 methanococ
33	80	6.6	411	1 PGK_METTH	O27121 methanobact

RESULT 1
ETXE_STAAU
ID ETXE_STAAU STANDARD; PRT; 257 AA.
AC P12993;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type E precursor (SEE).
GN ENTE.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-74.
RC STRAIN-MJB265;
RX MEDLINE=88257005; PubMed=3384800;
RA Couch J.L., Soltis M.T., Betley M.J.;
RT "Cloning and nucleotide sequence of the type E staphylococcal enterotoxin gene."
RL J. Bacteriol. 170:2954-2960(1988).
RN [2]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=96022387; PubMed=7552730;
RA Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
RT "Residues defining V beta specificity in staphylococcal enterotoxins."
RL Nat. Struct. Biol. 2:680-686(1995).
CC -!- CORACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary for the toxin interaction with MHC class II (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21319; AAA26617.1; .
CC PIR; A28179; A28179.
CC PDB; 1SEE; 15-OCT-95.
CC InterPro; IPR001961; Staph/Strep_toxin.
CC Pfam; PF01123; Staph_Strp_toxin; 1.
CC Pfam; PF02876; Staph_Strp_tox_C; 1.
CC PRINTS; PR00279; BACTRLTOXIN.
CC PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
CC PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
CC Enterotoxin; Toxin; Signal; Superautogen; Metal-binding; zinc;
KW 3D-structure; 1 27
FT SIGNAL 28 257 ENTEROTOXIN TYPE E.
FT CHAIN 28 257

34 80 6.6 1151 1 ITAL_HUMAN P56199 homo sapien
35 79.5 6.5 324 1 YB64_PASMU Q9CKY6 pasteurilla
36 79.5 6.5 637 1 VEL_HPV35 P27220 human papil
37 79.5 6.5 692 1 Y650_METJA O57852 methanococ
38 79 6.5 551 1 CHIT_NPVAC P41684 autographa
39 79 6.5 862 1 SLA2_BACAN P94217 bacillus an
40 79 6.5 3695 1 LMA5_HUMAN O15230 homo sapien
41 78.5 6.4 396 1 YJRI_YEAST P46992 saccharomyc
42 78.5 6.4 413 1 YGS0_YEAST P45819 saccharomyc
43 78.5 6.4 762 1 SLAP_ACEKI P22258 acetogenium
44 78 6.4 376 1 ADH1_ORYSA P20306 oryza sativ
45 78 6.4 607 1 DNK1_STRPN P95829 streptococ

```
FT METAL 211 211 ZINC (BY SIMILARITY).
FT METAL 249 249 ZINC (BY SIMILARITY).
FT METAL 251 251 ZINC (BY SIMILARITY).
SQ SEQUENCE 257 AA; 29358 MW; 27EDA94B97770CE3 CRC64;

Query Match 90.9%; Score 1107; DB 1; Length 257;
Best Local Similarity 89.7%; Pred. No. 2.2e-90;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFTNTLLFKGFFTG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFTNTLLFKGFFTG 84
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYYGQCAGGTPNKTACMYGGVTLHDNNRLT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYYGQCAGGTPNKTACMYGGVTLHDNNRLT 144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 EEKKVPINLWDGKQTTVPIDKVKTSKKEVTVOELDLQARHVLHGKFGLYNSDSFGGKVQ 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 145 EEKKVPINLWDGKQTTVPIDKVKTSKKEVTVOELDLQARHVLHGKFGLYNSDSFGGKVQ 204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 RGLIVFHSSEGSTVSYDLFDAGQYPTDLLRIYRDNKTISSTLSISLYLYTT 233
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 RGLIVFHSSEGSTVSYDLFDAGQYPTDLLRIYRDNKTISSTLSISLYLYTT 257

RESULT 2
ID ETVA_STAAM STANDARD; PRT; 257 AA.
AC P13163;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type A precursor (SEA).
GN ENTA OR MW1889.
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=2040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwana M., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RL acquired MRSA.";
RL Lancet 359:1819-1827(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FR1337;
RX MEDLINE=88086892; PubMed=3335483;
RA Betley M.J., Mekalanos J.J.;
RT "Nucleotide sequence of the type A staphylococcal enterotoxin gene.";
RL J. Bacteriol. 170:34-41(1988).
RN [3]
RP SEQUENCE OF 25-257.
RX MEDLINE=87222293; PubMed=3584106;
RA Huang I.-Y., Hughes J.L., Bergdoll M.S., Schantz E.J.;
RT "Complete amino acid sequence of staphylococcal enterotoxin A.";
RL J. Biol. Chem. 262:7006-7013(1987).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=95354648; PubMed=7628431;
RA Schad E.M., Zaitseva I., Zaitsev V.N., Dohlisten M., Kalland T.,
RA Schlievert P.M., Ohlendorf D.H., Svensson L.A.;
RT "Crystal structure of the superantigen staphylococcal enterotoxin
RL type A.";
RL EMBO J. 14:3292-3301(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=97113025; PubMed=8943278;
```

```
RA Sundstroem M., Hallen D., Svensson A., Schad E., Dohlisten M.,
RA Abrahamson L.;
RT "The Co-crystal structure of staphylococcal enterotoxin type A with
RT Zn2+ at 2.7-A resolution. Implications for major histocompatibility
RT complex class II binding.";
RL J. Biol. Chem. 271:32212-32216(1996).
RN [6]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=96022987; PubMed=7552730;
RA Swaminathan S., Furey W.F. Jr., Fletcher J., Sax M.;
RT "Residues defining V beta specificity in staphylococcal
RT enterotoxins.";
RL Nat. Struct. Biol. 2:680-686(1995).
RN [7]
RP COMPARISON OF STRUCTURE OF SEA AND SEC2.
RX MEDLINE=97334373; PubMed=9191070;
RA Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;
RT "A structural and functional comparison of staphylococcal
RT enterotoxins A and C2 reveals remarkable similarity and
RT dissimilarity.";
RL J. Mol. Biol. 269:270-280(1997).
CC -!- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
CC for the toxin interaction with MHC class II.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Secreted
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY A BACTERIOPHAGE.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP004828; BAB95754.1; -.
DR EMBL; M18970; AAA26681.1; -.
DR PIR; A28664; A28664.
DR PIR; A29566; A29566.
DR PDB; 1ESF; 11-JUL-96.
DR PDB; 1SEA; 15-OCT-95.
DR PDB; 1SXT; 19-NOV-97.
DR InterPro; IPR001961; Stap/Strep_toxin.
DR Pfam; PF01123; Stap_Strep_toxin; 1.
DR Pfam; PF02876; Stap_Strep_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;
KW 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 257 ENTEROTOXIN TYPE A.
FT DISULFID 120 130
FT METAL 211 211 ZINC.
FT METAL 249 249 ZINC.
FT METAL 251 251 ZINC.
FT CONFLICT 242 242 T -> S (IN REF. 3).
SQ SEQUENCE 257 AA; 29669 MW; ADEBF5BCALF14677 CRC64;

Query Match 77.8%; Score 948; DB 1; Length 257;
Best Local Similarity 76.4%; Pred. No. 2.3e-76;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFTNTLLFKGFFTG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFTNTLLFKGFFTG 84
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYYGQCAGGTPNKTACMYGGVTLHDNNRLT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AP003364; BAB58171.1; -
 DR EMBL; AP003135; BAB43097.1; -
 DR EMBL; X51661; CAA35972.1; -
 DR PIR; S11885; S11885
 DR PDB; 1JCK; 12-NOV-97
 DR InterPro; IPR001961; Stap/Strep_toxin.
 DR Pfam; PF01123; Stap_Strp_toxin; 1.
 DR Pfam; PF02876; Stap_Strp_tox_C; 1.
 DR PRINTS; PR00279; BACTRLTOXIN.
 DR PROSITE; PS00277; STAPH_STRP_TOXIN_1; 1.
 DR PROSITE; PS00278; STAPH_STRP_TOXIN_2; 1.
 DR Enterotoxin; Toxin; Signal; Superantigen; 3D-structure;
 KW Complete proteome.
 FT SIGNAL 1 27
 FT CHAIN 28 266
 FT DISULFID 120 137
 FT SEQUENCE 266 AA; 5ED8A32D11FCA59 CRC64;
 SQ

Query Match 23.4%; Score 284.5; DB 1; Length 266;
 Best Local Similarity 32.5%; Pred. No. 6.5e-18;
 Matches 76; Conservative 45; Mismatches 94; Indels 19; Gaps 9;

QY 11 DLKKSELOQTALGNLKOIYYNSKAITSS-KSADQFLNTLLPKGFTTGHVYNDLLV 69
 Db 37 DLKSSSEFTGT-MGNMK--YLYDDHYVSATKVKSDVKFLAHDLIYNSDKLKNYDKVKT 93
 QY 70 DLGSAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGCVTLHDNNRLTEE 122
 Db 94 ELLNEDLAKKYKDEVVDVYGSNYVNCYFSSKDNVGVGKTKMYGGITKHEGNHFDNG 153
 QY 123 --KKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDFSFGKVKV 180
 Db 154 NLQNVLRVY-ENKRNTISFE-VQTDKKSVAQELDIKARNFLINKNLYEENS--SPYE 209
 QY 181 RGLVFSHSEGSTVSYDLFDAQGYPD--TLRLRYRDNNTTSSLSLSLYLT 232
 Db 210 TGYIKFIENNGNTFWYDMMPAPGDKFQOSKYLMMYNDNKNKTVDKSKVIEVHLT 263

RESULT 5
 ETC2_STA00 STANDARD; PRT; 266 AA.
 AC P34071;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Enterotoxin type C-2 precursor (SEC2).
 GN EWC2.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxId=1280;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-66.
 RX MEDLINE=90277549; Pubmed=2543637;
 RA Bohach G.A., Schlievert P.M.;
 RT "Conservation of the biologically active portions of staphylococcal
 enterotoxins C1 and C2.";
 RL Infect. Immun. 57:2249-2252(1989).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=96027099; Pubmed=7582894;
 RA Papageorgiou A.C., Acharya K.R., Shapiro R., Passalacqua E.F.,
 RA Brehm R.D., Tranter H.S.;
 RT "Crystal structure of the superantigen enterotoxin C2 from
 Staphylococcus aureus reveals a zinc-binding site.";
 RL Structure 3:769-779(1995).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).

RX MEDLINE=96022987; Pubmed=7552730;
 RA Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
 RT "Residues defining V beta specificity in staphylococcal
 enterotoxins.";
 RL Nat. Struct. Biol. 2:680-686(1995).
 RN [4]
 RN COMPARISON OF STRUCTURE OF SEA AND SEC2.
 RX MEDLINE=97334373; Pubmed=9191070;
 RA Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;
 RT "A structural and functional comparison of staphylococcal
 enterotoxins A and C2 reveals remarkable similarity and
 dissimilarity.";
 RL J. Mol. Biol. 269:270-280(1997).
 CC -I- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
 for the toxin interaction with MHC class II.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
 STAPHYLOCOCCAL FOOD POISONING SYNDROME.
 CC -I- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
 FAMILY.
 DR PIR; A60114; A60114.
 DR PDB; 1STE; 23-DEC-96.
 DR PDB; 1SE2; 08-MAR-96.
 DR InterPro; IPR001961; Stap/Strep_toxin.
 DR Pfam; PF01123; Stap_Strp_toxin; 1.
 DR Pfam; PF02876; Stap_Strp_tox_C; 1.
 DR PRINTS; PR00279; BACTRLTOXIN.
 DR PROSITE; PS00277; STAPH_STRP_TOXIN_1; 1.
 DR PROSITE; PS00278; STAPH_STRP_TOXIN_2; 1.
 DR Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;
 KW 3D-structure.
 FT SIGNAL 1 27
 FT CHAIN 28 266
 FT DISULFID 120 137
 FT METAL 36 36
 FT METAL 110 110
 FT METAL 145 145
 FT METAL 149 149
 FT SEQUENCE 266 AA; 8407FB18536FAC08 CRC64;
 SQ

Query Match 22.8%; Score 277.5; DB 1; Length 266;
 Best Local Similarity 31.1%; Pred. No. 2.7e-17;
 Matches 76; Conservative 47; Mismatches 102; Indels 19; Gaps 9;

QY 1 SEKSEINEKDLRKSELOQTALGNLKOIYYNSKAITSS-KSADQFLNTLLPKGFTT 59
 Db 27 AESQPDPTDELKHSSEFTGT-MGNMK--YLYDDHYVSATKVKSDVKFLAHDLIYNSDK 83
 QY 60 GHPWYNDLLVDLSTAAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGCVT 112
 Db 84 KLKNDYKVKTELLNEDLAKKYKDEVVDVYGSNYVNCYFSSKDNVGVGKTKMYGGIT 143
 QY 113 LHDNNRLTEE--KKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLY 170.
 Db 144 KHGHNHFDNGNLQNVLRVY-ENKRNTISFE-VQTDKKSVAQELDIKARNFLINKNLY 201
 QY 171 NSDSFGGKQVORGLVHFHSSEGSTVSYDLFDAQGYPD--TLRLRYRDNNTTSSLSLSL 228
 Db 202 EFN5--SPYETGYIKFIENNGNTFWYDMMPAPGDKFQOSKYLMMYNDNKNKTVDKSKVIEV 259
 QY 229 YLYT 232
 Db 260 HLTT 263

RESULT 6
 SPEA_STRPY STANDARD; PRT; 251 AA.
 ID SPEA_STRPY
 AC P08095;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Exotoxin type A precursor (Scarlet fever toxin) (Erythrogenic toxin)


```

DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superantigen.
FT SIGNAL 1 27 ENTEROTOXIN TYPE C-1.
FT CHAIN 28 266
FT DISULFID 120 137
FT CONFLICT 177 177 D -> N (IN REF. 2).
SQ SEQUENCE 266 AA; 30546 MW; 3A7AB59A986853B CRC64;

Query Match 21.6%; Score 263.5; DB 1; Length 266;
Best Local Similarity 30.6%; Pred. No. 4.6e-16;
Matches 75; Conservative 45; Mismatches 104; Indels 21; Gaps 9;

Qy 1 SEKSEINEKDLRKSELOQTALGNLKOIY--YNSKAITSEKSDAQFLNTLFLKGF 58
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
27 AESQDPDPTDELHKASKFTG-LMENMKVLYDDHYVS---ATKVKSVDFLAHDLIYNISD 82
Qy 59 TGHPMYNLLVDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKACTACMGV 111
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
83 KKLKYNKVKTELLNEGAKYIKDEVDVGSNYVYVNFYSSKDNWGVKVGKTCMYGGI 142
Qy 112 TLHDNNRLTEE--KKVPINLWDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFL 169
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
143 TKHEGHFDNGNLQNLVRVY-ENKRNITSEF-VQTDKKSVTQAELDLKARNFLINKNL 200
Qy 170 YNSDSFGKVGORGLIVFHSSSGSVSYDLFDAQGYPD--TLRLIYRNTTISSTLSIS 227
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
201 YEFNS--SPYETGYIKFIENNGNTFWYDMPAPGDKFQSKYLMYNDKNTVDSKSVKIE 258
Qy 228 LYLTY 232
Db : : : : :
259 VHLTY 263

RESULT 9
ETXG_STAAM
ID ETXG_STAAM STANDARD; PRT; 258 AA.
AC 085382;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type G precursor (SEG).
GN ENTG OR SEG OR SAV1824 OR SA1642.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=FRI572;
RA Munson S.H., Tremaine M.T., Betley M.J., Welch R.A.;
RX MEDLINE=98298056; PubMed=9632603;
RT "Identification and characterization of staphylococcal enterotoxin
RT types G and I from Staphylococcus aureus.";
RL Infect. Immun. 66:3337-3348(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=MU50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cul L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yanashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.

```

```

CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; AF064773; AAC26660.1; --
DR EMBL; AP003363; BAB57986.1; --
DR EMBL; AP003135; BAB42910.1; --
DR HSSP; P01552; 1SBB.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_strp_toxin; 1.
DR Pfam; PF02876; Staph_strp_toxin_C; 1.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superantigen; Complete proteome.
FT SIGNAL 1 25
FT CHAIN 26 258 ENTEROTOXIN TYPE G.
FT DISULFID 116 133 BY SIMILARITY.
SQ SEQUENCE 258 AA; 29940 MW; E2982101701D012C CRC64;

Query Match 19.8%; Score 241; DB 1; Length 258;
Best Local Similarity 28.8%; Pred. No. 4.3e-14;
Matches 72; Conservative 48; Mismatches 88; Indels 42; Gaps 12;

Qy 3 KSEINEKDLRKSELOQTALGNLKOIY--YNSKAITSEKSDAQFLNTLFLKGF 59
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
30 KLELNKVSQYKNN--KGT-MGNVMNLVTPPVVEGRGVINSR---QFLSHDLIEP--I 79
Qy 60 GHPWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQCA-----GGTPNKACTAC 107
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
80 EYKSYNEVKTELENTLENNYKDKKVDIFGVFYTCIIPKSEPDINQFEG--CCM 134
Qy 108 YGGVTLH--DNNRLTEKKVPINLWDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHG 165
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
135 YGGLTFNSENER---DKLITVQVTDNRQSLG--FTITNNKMWVTIQELDYKARHWT 189
Qy 166 KGLYNSDSFGKVGORGLIVFHSSSGSVSYDLFDAQGYPD--LRLIYRNTTISST 222
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
190 EKKLYEFD--GSAFESGYIKFTKNTSFWDLFPKELVPVPVKFLNIYGDKNVDSK 247
Qy 223 SLSISLYLT 232
Db : : : : :
248 SIKMEVFLNT 257

RESULT 10
SPEH_STRPY
ID SPEH_STRPY STANDARD; PRT; 236 AA.
AC Q9X5C8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Exotoxin type H precursor (SPE H).
GN SPEH OR SPH1008.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M15;
RX MEDLINE=99093428; PubMed=9874566;
RA Proft T., Mofatt S.L., Berkahn C.J., Fraser J.D.;
RT "Identification and characterization of novel superantigens from
RT Streptococcus pyogenes.";
RL J. Exp. Med. 189:89-102(1999).
RN [2]

```

```
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001)
CC -I- SUBUNIT: Binds to major histocompatibility complex class II beta
CC chain.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- DISEASE: Mitogenic for human peripheral blood lymphocytes.
CC -I- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF124500; AAD30989.1; -.
CC EMBL: AE006546; AAK33907.1; -.
CC HSP: P13163; 1SX7.
CC InterPro: IPR001961; Staph/Strep_toxin.
CC Pfam: PF01123; Staph_Strp_toxin; 1.
CC Pfam: PF02876; Staph_Strp_tox_C; 1.
CC PROSITE: PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
CC PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
CC Toxin; Signal; Complete proteome.
CC SIGNAL 1 32 POTENTIAL.
CC CHAIN 33 236 EXOTOXIN TYPE H.
CC SEQUENCE 236 AA; 27485 MW; 16352923907AD40D CRC64;
Query Match 15.28; Score 185; DB 1; Length 236;
Best Local Similarity 26.78; Pred. No. 3.3e-09;
Matches 56; Conservative 46; Mismatches 84; Indels 24; Gaps 9;
Qy 25 NLKQIYYNSKAI-TSEKSDAQFLNTLLFKGFTGHPWYN-----DLLVDLGSTAAATSE 79
Db 42 NLESLYKHDLSLEADSTKNSPDIVTSHL-----KYSVKDKNLSVFFERDWSIQE 92
Qy 80 YEGSSVDLYGAYGVQCAGGTPNKACMGVTLHDNNRLTEKKVPINLWIDGKQTTVP 139
Db 93 FKDEVDIYALSQEVCE--CPGKRYEAFGGITLTNSEK--KEIKVPVNWWDKSKQ--P 146
Qy 140 IDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGKVGQGLIVFHSSEGSTVSYDL- 198
Db 147 PMFITVKNPKVTAQEVDIKVRKLLIKKYDIYNNRE--QKYSKGTVTLDLNSGKDIVFDLY 204
Qy 199 FDAQGYPTDLLRIYRDNTTISSTLSLSL 228
Db 205 YFGNGDF--NSMLKIYSNNERIDSTQFHVYD 233
RESULT 11
SPEC_STRPY
ID SPEC_STRPY STANDARD; PRT; 234 AA.
AC Q9X5C7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exotoxin type G precursor (SPE G).
GN SPEG OR SPY0212.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=M1;
RX MEDLINE=99093428; PubMed=9874566;
RA Proft T., Moffatt S.L., Berkahn C.J., Fraser J.D.;
RT "Identification and characterization of novel superantigens from
RL Streptococcus pyogenes."
RN J. Exp. Med. 189:89-102(1999).
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001)
CC -I- SUBUNIT: Binds to major histocompatibility complex class II beta
CC chain.
CC -I- DISEASE: Mitogenic for human peripheral blood lymphocytes.
CC -I- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF124499; AAD30988.1; -.
CC EMBL: AE006489; AAK33303.1; -.
CC HSP: P13380; 1AN8.
CC InterPro: IPR001961; Staph/Strep_toxin.
CC Pfam: PF01123; Staph_Strp_toxin; 1.
CC Pfam: PF02876; Staph_Strp_tox_C; 1.
CC PRINTS: PR00279; BACTRLTOXIN.
CC PROSITE: PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
CC PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
CC Toxin; Signal; Complete proteome.
CC SIGNAL 1 24 POTENTIAL.
CC CHAIN 25 234 EXOTOXIN TYPE G.
CC SEQUENCE 234 AA; 27262 MW; 49525C49E4BA2052 CRC64;
Query Match 15.18; Score 183.5; DB 1; Length 234;
Best Local Similarity 27.68; Pred. No. 4.4e-09;
Matches 56; Conservative 37; Mismatches 81; Indels 29; Gaps 7;
Qy 33 NSKAITSEKSDAQFLNTLLFKGFTGHPWYNLLVDLGSTAAATSEYEGSSVDLYGAY 92
Db 57 NSIHNTKQRSECI-----YVDSIVSLGIT--DQFIKGDKVDVFGLPY 99
Qy 93 GYCAGGTPNKACMGVTLHDNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKKEVTV 152
Db 100 NF-----SPPYVDNIYGGIVKHSNQGKSLQFVGTLNQDGKETYLPSEAVRIKKQFTL 153
Qy 153 QELDLQARHYLHGKFGLYNSDSFGKVGQGLIVFHSSEGSTVSYDLFDAQGY--PDFLL 210
Db 154 QEFDFKTKFLMEKYNIDSESE---RYTSGSLFLATKDSKHYEVDLFPNKDKLLSRDSFF 210
Qy 211 RIYRDNTTISSTLS-LSLYLT 232
Db 211 KRYKDNKIFNSEEISHFDIYLT 233
RESULT 12
SPEC_STRPY
ID SPEC_STRPY STANDARD; PRT; 235 AA.
AC P13380;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exotoxin type C precursor (SPE C).
```



```
RT "Refined structures of three crystal forms of toxic shock syndrome
RL toxin-1 and of a tetramutant with reduced activity.";
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS) OF MUTANTS.
RX MEDLINE=98254504; PubMed=9585531;
RA Earhart C.A., Mitchell D.T., Murray D.L., Pinheiro D.M., Matsumura M.,
RA Schlievert P.M., Ohlendorf D.H.;
RT "Structures of five mutants of toxic shock syndrome toxin-1 with
RT reduced biological activity.";
RL Biochemistry 37:7194-7202(1998).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: THIS TOXIN IS RESPONSIBLE FOR THE SYMPTOMS OF TOXIC
CC SHOCK SYNDROME.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02615; AAA26682.1; .
DR PIR: A24606; XCSAS1.
DR PDB: 2TSS; 24-DEC-97.
DR PDB: 3TSS; 24-DEC-97.
DR PDB: 4TSS; 24-DEC-97.
DR PDB: 5TSS; 24-DEC-97.
DR PDB: 10IL; 12-AUG-97.
DR PDB: 20IL; 12-AUG-97.
DR PDB: 1A07; 18-NOV-98.
DR PDB: 1TS2; 16-DEC-98.
DR PDB: 1TS3; 16-DEC-98.
DR PDB: 1TS4; 16-DEC-98.
DR PDB: 1TS5; 16-DEC-98.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR InterPro: IPR003999; Staph_toxin.
DR Pfam: PF01123; Staph_strep_toxin; 1.
DR Pfam: PF02876; Staph_strep_tox_C; 1.
DR PRINTS: PR01501; TOXICSTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
KW Toxin; Superantigen; Signal; 3D-structure.
FT SIGNAL 1 40
FT CHAIN 41 234
SQ SEQUENCE 234 AA; 26306 MW; E95789FF9A1D7AB4 CRC64;

Query Match 8.2%; Score 100; DB 1; Length 234;
Best Local Similarity 24.5%; Pred.No. 0.099;
Matches 48; Conservative 33; Mismatches 61; Indels 54; Gaps 12;
QY 25 NLKQI---YYNKAITSSE-----KSAQFLNTLLFKGFTGHPWYNDLL-- 68
DB 45 NIKDLLWYSSGSDFTNSEVLDNSLGNRIKNTDGSII-SLIIFPS-----PYSPATK 98
QY 69 ----VDLGSTAATSEYGGSDVLYGAYGQCAGGTPNKTACMYGGVTLHDNNRLTEKKV 125
DB 99 GEKVLD-NTKRKKSQHTS---EGTYIHFIQISGVT-----NTEKLPPTIEL 140
QY 126 PINLWIDGKQTVPIDKVTKSKKVTQVQLDQARHYLHGKFGLY-NSDSFGG--KVQRG 182
DB 141 PLKRVKHGKDSPLKGYG-KFKDKQLAISTLDPEIRHQLTQHGLYRSSDKTGKGYKIT-- 197
QY 183 LIVFHSSEGSTVSVDL 198
DB 198 -----MNDGSTYQSDL 208

RESULT 14
RI9E_THEVO
```

```
ID RI9E_THEVO STANDARD; PRT; 150 AA.
AC Q97CU4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S19E.
GN RPS19E OR TV0007 OR TVG0008143.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS51 / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kunoshima T., Yamamoto Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Nunoshima T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of Thermoplasma volcanium";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
CC -!- SIMILARITY: BELONGS TO THE S19E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000991; BAB59149.1; .
DR InterPro: IPR001266; Ribosomal_S19E.
DR Pfam: PF01090; Ribosomal_S19e; 1.
DR PROSITE: PS00628; RIBOSOMAL_S19E; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 150 AA; 17063 MW; FD5881CF684EB415 CRC64;

Query Match 7.8%; Score 95.5; DB 1; Length 150;
Best Local Similarity 29.5%; Pred.No. 0.14;
Matches 23; Conservative 17; Mismatches 37; Indels 1; Gaps 1;
QY 116 NNRLTEERKKVPINLWIDGKQTVPIDKVTKSKKVTQVQLDQARHYLHGKFGLYN-SDS 174
DB 17 SGKLEKKIKPEPNVFKVTKGVSKEKPPLODDWIVVRAASMLRKLINGYLIGISMSE 76
QY 175 FGGKVORGLVIFHSSEGS 192
DB 77 YGGKVDGRGSKRYHAASGS 94

RESULT 15
PI15_MYCHR
ID PI15_MYCHR STANDARD; PRT; 979 AA.
AC P41508;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PI15 protein.
OS Mycoplasma hyorhinis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2100;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91138990; PubMed=1825306;
RA Notaricola S.M., McIntosh M.A., Wise K.S.;
RT "A Mycoplasma hyorhinis protein with sequence similarities to
RT nucleotide-binding enzymes.";
RL Gene 97:77-85(1991).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS
CC FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYCOPLASMA
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC P115.
CC -----
CC EMBL; M34956; AAA25423.1; -
CC PIR; JQ0894; JQ0894.
CC InterPro: IPR003439; ABC_transportr.
CC InterPro: IPR005289; GTP-binding_dom.
CC InterPro: IPR003405; SMC_C.
CC InterPro: IPR003395; SMC_N.
CC Pfam; PF02463; SMC_N; 1.
CC Pfam; PF02483; SMC_C; 1.
CC TIGRfams; TIGR00650; MG442; 2.
KW ATP-binding; Coiled coil.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT DOMAIN 169 224 COILED COIL (POTENTIAL).
FT DOMAIN 231 400 COILED COIL (POTENTIAL).
FT DOMAIN 569 821 COILED COIL (POTENTIAL).
FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).
SQ SEQUENCE 979 AA; 110566 MW; 30D51C5B56280F4 CRC64;

Query Match 7.6%; Score 93; DB 1; Length 979;
Best Local Similarity 23.3%; Pred. No. 2.4;
Matches 55; Conservative 41; Mismatches 92; Indels 48; Gaps 9;

QY 17 ELQGTALGNLKOIYYNSKAITSSKSAQO-----FLTNLLFKGFFTGH 61
DB 17 ELQGTALGNLKOIYYNSKAITSSKSAQO-----FLTNLLFKGFFTGH 61
QY 62 PWNDLLV---DLGSTAANTSEYEGSSVDLYGAYGYQCAGGTPNKTACHY---GGVTLHD 115
DB 62 PWNDLLV---DLGSTAANTSEYEGSSVDLYGAYGYQCAGGTPNKTACHY---GGVTLHD 115
QY 113 YFNDOPVRYKDKLNKAVESGSKSLAISQGTISEIAEATPEQRKAVIEEAAGTSKYK 172
DB 113 YFNDOPVRYKDKLNKAVESGSKSLAISQGTISEIAEATPEQRKAVIEEAAGTSKYK 172
QY 116 NNRLTEBKVPINLWIDGKQTTVPIDKVTSSKKEV--TVQELDLQ---ARHYLHGKFGLY 170
DB 116 NNRLTEBKVPINLWIDGKQTTVPIDKVTSSKKEV--TVQELDLQ---ARHYLHGKFGLY 170
QY 173 LDKEEAQKL-----IFTNDAIDKQAIKELERQVNSLDKQASKAKIYLEKSKALE 224
DB 173 LDKEEAQKL-----IFTNDAIDKQAIKELERQVNSLDKQASKAKIYLEKSKALE 224
QY 171 NSDSFGKGVQRLIV-----FHSSEGTVSVDLFDAGQVPTDLLRIYRONTTISST 222
DB 171 NSDSFGKGVQRLIV-----FHSSEGTVSVDLFDAGQVPTDLLRIYRONTTISST 222
QY 225 S-----VEVGLIVDLNLFNEKLNNTSLLEVEQQRNDLELNTQTVESISSQT 273
DB 225 S-----VEVGLIVDLNLFNEKLNNTSLLEVEQQRNDLELNTQTVESISSQT 273

RESULT 16
DPOL_METJA STANDARD; PRT; 1634 AA.
AC 058295;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) [Contains: Mja pol-1 intein; Mja pol-2 intein].
GN POL OR MJ0885.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcaceae;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
```

```
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL Science 273:1058-1073(1996).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + (DNA)(N).
CC -!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67532; AAB98889.1; -
CC HSSP; P56689; 1TGO.
CC TIGR; MJ0885.
CC InterPro: IPR002054; DNA_pol_B.
CC InterPro: IPR003586; Hedgehog_hintc.
CC InterPro: IPR003587; Hedgehog_hintn.
CC InterPro: IPR002203; Intein.
CC InterPro: IPR004042; Intein_endonuc.
CC InterPro: IPR004578; Pol2.
CC Pfam; PF00136; DNA_pol_B; 3.
CC Pfam; PF03104; DNA_pol_B_exo; 1.
CC PRINTS; P00106; DNAPOLB.
CC SMART; SM00305; Hintc; 2.
CC SMART; SM00306; Hintn; 2.
CC SMART; SM00486; POLBC; 1.
CC TIGRfams; TIGR00592; pol2; 1.
CC PROSITE; PS00116; DNA_POLYMERASE_B; 1.
CC PROSITE; PS00818; INTEIN_C_TER; 2.
CC PROSITE; PS00819; INTEIN_ENDONUCLEASE; 2.
CC PROSITE; PS00817; INTEIN_N_TER; 2.
CC Transfaser; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Autocatalytic cleavage; Protein splicing;
KW Complete proteome.
FT CHAIN 1 425 POL, 1ST PART (POTENTIAL).
FT CHAIN 426 794 MJA POL-1 INTEIN (POTENTIAL).
FT CHAIN 795 882 POL, 2ND PART (POTENTIAL).
FT CHAIN 883 1358 MJA POL-2 INTEIN (POTENTIAL).
FT CHAIN 1359 1634 POL, 3RD PART (POTENTIAL).
SQ SEQUENCE 1634 AA; 191708 MW; 84A1FAFAB1F97DDD CRC64;

Query Match 7.5%; Score 91; DB 1; Length 1634;
Best Local Similarity 21.6%; Pred. No. 6.8;
Matches 50; Conservative 36; Mismatches 79; Indels 66; Gaps 10;

QY 12 LRKSELOGTALGNLKOIYYNSKAITSSKSAQDLFTNLLFKGFFTGHPWNLLVDL 71
DB 12 LRKSELOGTALGNLKOIYYNSKAITSSKSAQDLFTNLLFKGFFTGHPWNLLVDL 71
QY 614 VRRKKGTKAITLGCARKDIYKIEIL---KNKEKYLPAAIL-RGFEFGDGYVNTVRRV 669
DB 614 VRRKKGTKAITLGCARKDIYKIEIL---KNKEKYLPAAIL-RGFEFGDGYVNTVRRV 669
QY 72 GSTAATSEYE----GSSVDLYGAYGYQCAGGTPNKTACMYGVTLHONNRLTEKKVP 126
DB 72 GSTAATSEYE----GSSVDLYGAYGYQCAGGTPNKTACMYGVTLHONNRLTEKKVP 126
QY 670 VYNOGTNNYDKIKFIASLLDLRLGIKYSF-----YTSYEERKKLKRKY 713
DB 670 VYNOGTNNYDKIKFIASLLDLRLGIKYSF-----YTSYEERKKLKRKY 713
QY 127 INLWIDGKQTTVPIDKVKTS-----KKEVTVQELDLQARHYLHGKFGLYNSD----- 173
DB 127 INLWIDGKQTTVPIDKVKTS-----KKEVTVQELDLQARHYLHGKFGLYNSD----- 173
QY 714 IEIFSGK-----DLIKFSILISFISRRKNLLNEIRKTKLYKIGDYGYDLDQCVS 766
DB 714 IEIFSGK-----DLIKFSILISFISRRKNLLNEIRKTKLYKIGDYGYDLDQCVS 766
QY 174 --SFGKGV-----QRGLIVFHSSEGS--TVSYDLFDAGQVPTDL 209
DB 174 --SFGKGV-----QRGLIVFHSSEGS--TVSYDLFDAGQVPTDL 209
QY 767 LESYKGEVYDLTLEGRPPYFANGILTHNSLYPSIIISYNI-----SPDTL 811
DB 767 LESYKGEVYDLTLEGRPPYFANGILTHNSLYPSIIISYNI-----SPDTL 811

RESULT 17
ID R19E_THEAC STANDARD; PRT; 150 AA.
```

AC Q9HM21;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 30S ribosomal protein S19E.
 DE 30S ribosomal protein S19E.
 GN RPS19E OR TA0050.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Meves H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermoacidophilic scavenger thermoplasma
 acidophilum.";
 RL Nature 407:508-513(2000).
 CC -!- SIMILARITY: BELONGS TO THE S19E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL445063; CAC11198.1;
 DR InterPro; IPR001266; Ribosomal_S19E.
 DR Pfam; PF01090; Ribosomal_S19e; 1.
 DR ProDom; PD003854; Ribosomal_S19e; 1.
 DR PROSITE; PS00628; RIBOSOMAL_S19E; FALSE_NEG.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 150 AA; 17109 MW; F3215035CE0DC22C CRC64;
 Query Match 7.3%; Score 88.5; DB 1; Length 150;
 Best Local Similarity 26.9%; Pred. No. 0.59;
 Matches 21; Conservative 20; Mismatches 36; Indels 1; Gaps 1;
 QY 116 NNRLTEKVPNLTWDGQTTVPIDKVTSKKEVTVQELDQARHYLHGKFGLYN-SDS 174
 DB 17 SEKLSEKKIAEPDQSKYKVTGISEKSPVNRDIYVRAAAMLRKLYINGILGISMSE 76
 QY 175 FGGKVORGLIVFHSSRGS 192
 DB 77 YGKVDGRGSKRTHAAGS 94
 RESULT 18
 ID TYPH_MYCPI STANDARD; PRT; 419 AA.
 AC P4717;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Thymidine phosphorylase (EC 2.4.2.4) (TDRPASE).
 GN DEOA.
 OS Mycoplasma pirum.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2122;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BER.
 RX MEDLINE=93352438; PubMed=8349569;
 RA Tham T.N., Ferris S., Kovacic R., Montagnier L., Blanchard A.;
 RT "Identification of Mycoplasma pirum genes involved in the salvage
 RT pathways for nucleosides."
 RL J. Bacteriol. 175:5281-5285(1993).
 CC -!- FUNCTION: THE ENZYMES WHICH CATALYZE THE REVERSIBLE PHOSPHORYLOSIS
 CC OF PYRIMIDINE NUCLEOSIDES ARE INVOLVED IN THE DEGRADATION OF THESE
 CC COMPOUNDS AND IN THEIR UTILIZATION AS CARBON AND ENERGY SOURCES.

CC OR IN THE RESCUE OF PYRIMIDINE BASES FOR NUCLEOTIDE SYNTHESIS.
 CC -!- CATALYTIC ACTIVITY: Thymidine + phosphate = thymine + 2-deoxy-D-
 CC ribose 1-phosphate.
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE THYMIDINE/PYRIMIDINE-NUCLEOSIDE
 CC PHOSPHORYLASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L13289; AAA25432.1;
 DR HSP; P77836; IBRW.
 DR InterPro; IPR000312; Glycos_transf_3.
 DR InterPro; IPR000053; Thymid_phosphils.
 DR Pfam; PF00591; Glycos_transf_3; 1.
 DR Pfam; PF02885; Glycos_trans_3N; 1.
 DR ProDom; PD005916; Thymid_phosphils; 1.
 DR PROSITE; PS00647; THYMID_PHOSPHORYLASE; 1.
 KW Transferase; Glycosyltransferase.
 SQ SEQUENCE 419 AA; 47278 MW; C1620F2EF33929E3 CRC64;
 Query Match 7.2%; Score 88; DB 1; Length 419;
 Best Local Similarity 27.0%; Pred. No. 2.3;
 Matches 38; Conservative 22; Mismatches 57; Indels 24; Gaps 7;
 QY 47 FLTNTLLFKG-FFTGHPWYNLLVLGSLTAATSE-----YEGSSVDLXGAYKG 93
 DB 57 FLTKAMIDSGKIYKHPPEYKKILIDKHSGGIGDKVSIARPLVSLFGLVAKLSGRGLG 116
 QY 94 YOCAGTTPNKTACMYGGVTLHDNNRLTEKKV--PINLMIDGK-QTTVPIDKVKTSKKEV 150
 DB 117 F--TGTIDKLE----SINVNTDIDLKSKILNIANMFIVGQTNIDIVPADKLLYALRDV 170
 QY 151 --TVQELDQARHYLHGKFG 169
 DB 171 TGTVDLSPLIAAISLSKKPAL 191
 RESULT 19
 ID HRP3_SCHPO STANDARD; PRT; 1388 AA.
 AC O14139;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chromodomain helicase hrp3.
 GN HRP3 OR SPAC3G6.01.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gencies S., Gobbe A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,


```
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X63152; CAA44855.1; -.
EMBL; Z14065; CAA78448.1; -.
EMBL; AE004529; AAG04347.1; -.
PIR; S25008; S25008.
PIR; S23771; S23771.
MEROPS; S43.001; -.
InterPro: IPR005318; Peptidase_S43.
PFam; PF03573; Peptidase_S43; 1.
KW Hydrolase; Serine protease; Transport; Outer membrane; Transmembrane;
KW Porin; Signal; Complete proteome.
FT SIGNAL 1 23
FT CHAIN 24 443
FT ACT_SITE 179 179
FT ACT_SITE 231 231
FT ACT_SITE 319 319
FT MUTAGEN 179 179
FT MUTAGEN 231 231
FT MUTAGEN 319 319
FT MUTAGEN 390 390
FT CONFLICT 44 44
FT SEQUENCE 443 AA; 48360 MW; E083FPE074DCF864 CRC64;
-----
Query Match 6.98; Score 84.5; DB 1; Length 443;
Best Local Similarity 23.28; Pred. No. 5.1;
Matches 60; Conservative 30; Mismatches 98; Indels 71; Gaps 14;
QY 15 KSELGTLGNLKIYVNSKAITSSKSAQDLTNTLLFGFTGHPWYNDLLVDLGLST 74
DB 34 KGFIEDSSLDLLRNYFNROCKSGSRVD-----WTQGFLLTY-----ESFT 78
QY 75 AATSEYEGSSVDLYGAYGYOCAGTGNKTA-----CMYGGVTLHNNRLTEKKVPI 128
DB 79 QGT---VGFQVDAFG-VLGLK-LDGTSDKTGNLPVNDGKPRDYSRAGAVKVRISK 133
QY 129 ---LWIDGKQTTVPIDKVKTSK-----KEVTQVQLDLOARHYLHGK----- 166
DB 134 TMLKWE-MQTPAPFAAGGSKLFPQTATGQLQSSEFEGLDLEAGHTEGKPTTVKSR 192
QY 167 ---FGLY-----NSDSFGKVQRLIVFHSSEGSTVSYDLFDAGQGYPTLLRIYRNT 217
DB 193 GELYATYAGETAKSADFGR-----YATDNLASLYCAELEDIYRQY-----YLSN 242
QY 218 TI---SSTSLSLSLYLYTT 233
DB 243 TIPLASDQSLGDFDFNIYRT 261
RESULT 21
DNAX_STRMU STANDARD; PRT; 590 AA.
AC O06942;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa
protein) (HSP70) (Fragment).
GN DNAX.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
-----
RN RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=97426041; PubMed=9282745;
RA Jayaraman G.C., Penders J.E., Burne R.A.;
RT "Transcriptional analysis of the Streptococcus mutans hrcA, grpE and
dnak genes and regulation of expression in response to heat shock and
environmental acidification.";
RL Mol. Microbiol. 25:329-341(1997).
CC -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC -!- INDUCTION: BY HEAT SHOCK.
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U78296; AAC45612.1; -.
HSSP; P04475; 2BPR.
InterPro: IPR001023; Hsp70.
PFam; PF00012; HSP70; 1.
PRODOM; PD000089; Hsp70; 1.
PROSITE; PS00297; HSP70_1; 1.
PROSITE; PS00329; HSP70_2; 1.
PROSITE; PS01036; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
FT MOD_RES 173 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT NON_TER 590 590
FT SEQUENCE 590 AA; 63258 MW; 90787ADF08F702B9 CRC64;
-----
Query Match 6.98; Score 84; DB 1; Length 590;
Best Local Similarity 27.78; Pred. No. 8;
Matches 38; Conservative 18; Mismatches 63; Indels 18; Gaps 5;
QY 69 VDLGST-AATSEYEGSSVDLYGAYGYOCAGTGNKTA-CMYGGVTLHNNRLTEKKVPI 127
DB 7 IDLGTNSAVAVLEGTESKIITANPEGNT---TPSVVSKNGEIIIVGDAARQAVNPET 63
QY 128 NLWIDGKQTTVPIDKVKTSKKEVTQVQLDLOARHYLHGKFGLYNSDSFGKVGQRLIVFH 187
DB 64 ILSIKSMGT--SEKVSANAKETPQETISAMILQYLGK----YAEDYLGEKVEKAVI--- 114
QY 188 SSEGSTVSYDLFDAGQ 204
DB 115 -----TVPAYFNDAQRQ 126
RESULT 22
ITAS_XENLA
ID ITA5_XENLA STANDARD; PRT; 1050 AA.
AC Q06274;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integrin alpha-5 precursor (Fibronectin receptor alpha subunit)
DE (Integrin alpha-5) (VLA-5).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9534494; PubMed=7619730;
RA Joos T.O., Whittaker C.A., Meng F., Desimone D.W., Gnau V.,
RA Hausen P.;
RT "Integrin alpha 5 during early development of Xenopus laevis.";
RL Mech. Dev. 50:187-199(1995).
RN [2]
```

```
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X63152; CAA44855.1; -.
EMBL; Z14065; CAA78448.1; -.
EMBL; AE004529; AAG04347.1; -.
PIR; S25008; S25008.
PIR; S23771; S23771.
MEROPS; S43.001; -.
InterPro: IPR005318; Peptidase_S43.
PFam; PF03573; Peptidase_S43; 1.
KW Hydrolase; Serine protease; Transport; Outer membrane; Transmembrane;
KW Porin; Signal; Complete proteome.
FT SIGNAL 1 23
FT CHAIN 24 443
FT ACT_SITE 179 179
FT ACT_SITE 231 231
FT ACT_SITE 319 319
FT MUTAGEN 179 179
FT MUTAGEN 231 231
FT MUTAGEN 319 319
FT MUTAGEN 390 390
FT CONFLICT 44 44
FT SEQUENCE 443 AA; 48360 MW; E083FPE074DCF864 CRC64;
-----
Query Match 6.98; Score 84.5; DB 1; Length 443;
Best Local Similarity 23.28; Pred. No. 5.1;
Matches 60; Conservative 30; Mismatches 98; Indels 71; Gaps 14;
QY 15 KSELGTLGNLKIYVNSKAITSSKSAQDLTNTLLFGFTGHPWYNDLLVDLGLST 74
DB 34 KGFIEDSSLDLLRNYFNROCKSGSRVD-----WTQGFLLTY-----ESFT 78
QY 75 AATSEYEGSSVDLYGAYGYOCAGTGNKTA-----CMYGGVTLHNNRLTEKKVPI 128
DB 79 QGT---VGFQVDAFG-VLGLK-LDGTSDKTGNLPVNDGKPRDYSRAGAVKVRISK 133
QY 129 ---LWIDGKQTTVPIDKVKTSK-----KEVTQVQLDLOARHYLHGK----- 166
DB 134 TMLKWE-MQTPAPFAAGGSKLFPQTATGQLQSSEFEGLDLEAGHTEGKPTTVKSR 192
QY 167 ---FGLY-----NSDSFGKVQRLIVFHSSEGSTVSYDLFDAGQGYPTLLRIYRNT 217
DB 193 GELYATYAGETAKSADFGR-----YATDNLASLYCAELEDIYRQY-----YLSN 242
QY 218 TI---SSTSLSLSLYLYTT 233
DB 243 TIPLASDQSLGDFDFNIYRT 261
RESULT 21
DNAX_STRMU STANDARD; PRT; 590 AA.
AC O06942;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa
protein) (HSP70) (Fragment).
GN DNAX.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
-----
RN RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=97426041; PubMed=9282745;
RA Jayaraman G.C., Penders J.E., Burne R.A.;
RT "Transcriptional analysis of the Streptococcus mutans hrcA, grpE and
dnak genes and regulation of expression in response to heat shock and
environmental acidification.";
RL Mol. Microbiol. 25:329-341(1997).
CC -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC -!- INDUCTION: BY HEAT SHOCK.
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U78296; AAC45612.1; -.
HSSP; P04475; 2BPR.
InterPro: IPR001023; Hsp70.
PFam; PF00012; HSP70; 1.
PRODOM; PD000089; Hsp70; 1.
PROSITE; PS00297; HSP70_1; 1.
PROSITE; PS00329; HSP70_2; 1.
PROSITE; PS01036; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
FT MOD_RES 173 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT NON_TER 590 590
FT SEQUENCE 590 AA; 63258 MW; 90787ADF08F702B9 CRC64;
-----
Query Match 6.98; Score 84; DB 1; Length 590;
Best Local Similarity 27.78; Pred. No. 8;
Matches 38; Conservative 18; Mismatches 63; Indels 18; Gaps 5;
QY 69 VDLGST-AATSEYEGSSVDLYGAYGYOCAGTGNKTA-CMYGGVTLHNNRLTEKKVPI 127
DB 7 IDLGTNSAVAVLEGTESKIITANPEGNT---TPSVVSKNGEIIIVGDAARQAVNPET 63
QY 128 NLWIDGKQTTVPIDKVKTSKKEVTQVQLDLOARHYLHGKFGLYNSDSFGKVGQRLIVFH 187
DB 64 ILSIKSMGT--SEKVSANAKETPQETISAMILQYLGK----YAEDYLGEKVEKAVI--- 114
QY 188 SSEGSTVSYDLFDAGQ 204
DB 115 -----TVPAYFNDAQRQ 126
RESULT 22
ITAS_XENLA
ID ITA5_XENLA STANDARD; PRT; 1050 AA.
AC Q06274;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integrin alpha-5 precursor (Fibronectin receptor alpha subunit)
DE (Integrin alpha-5) (VLA-5).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9534494; PubMed=7619730;
RA Joos T.O., Whittaker C.A., Meng F., Desimone D.W., Gnau V.,
RA Hausen P.;
RT "Integrin alpha 5 during early development of Xenopus laevis.";
RL Mech. Dev. 50:187-199(1995).
RN [2]
```

```

SEQUENCE OF 318-393 FROM N.A.
MEDLINE=94008528; PubMed=8404528;
Whittaker C.A., Desimone D.W.;
"Integrin alpha subunit mRNAs are differentially expressed in early
Xenopus embryos";
Development 117:1239-1249(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRONECTIN.
CC IT RECOGNIZES THE SEQUENCE R-G-D IN ITS LIGANDS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
CC DISULFIDE BOND. ALPHA-5 ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; U12683; AAA99668.1; -
DR EMBL; LI0191; AAA16249.1; -
DR HSSP; P06756; IJY2.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF01839; FG-GAP; 5.
DR PRINTS; PRO1185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 5.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 1050 INTEGRIN ALPHA-5.
FT CHAIN 33 932 INTEGRIN ALPHA-5 HEAVY CHAIN (POTENTIAL).
FT CHAIN 933 1050 INTEGRIN ALPHA-5 LIGHT CHAIN (POTENTIAL).
FT DOMAIN 33 996 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 997 1022 POTENTIAL.
FT DOMAIN 1023 1050 CYTOPLASMIC (POTENTIAL).
FT REPEAT 48 110 FG-GAP 1.
FT REPEAT 120 188 FG-GAP 2.
FT REPEAT 189 246 FG-GAP 3.
FT REPEAT 259 312 FG-GAP 4.
FT REPEAT 313 373 FG-GAP 5.
FT REPEAT 379 436 FG-GAP 6.
FT REPEAT 442 494 FG-GAP 7.
FT CA_BIND 324 332 POTENTIAL.
FT CA_BIND 390 398 POTENTIAL.
FT CA_BIND 454 462 POTENTIAL.
FT SITE 1025 1029 GFGR MOTIF.
FT DISULFID 90 99 BY SIMILARITY.
FT DISULFID 145 166 BY SIMILARITY.
FT DISULFID 182 195 BY SIMILARITY.
FT DISULFID 502 513 BY SIMILARITY.
FT DISULFID 519 575 BY SIMILARITY.
FT DISULFID 636 642 BY SIMILARITY.
FT DISULFID 708 721 BY SIMILARITY.
FT DISULFID 862 910 INTERCHAIN (BY SIMILARITY).
FT DISULFID 917 922 BY SIMILARITY.
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 521 521 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 600 600 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 6.9%; Score 84; DB 1; Length 1050;
Best Local Similarity 21.9%; Pred. No. 16;
Matches 60; Conservative 38; Mismatches 94; Indels 82; Gaps 14;

FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 763 763 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 861 861 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1050 AA; 115961 MW; 10ED961535B8D918 CRC64;

QY 12 LRKKSELO-GTALGNLKKQIYYNSKAITS-SEKSAQOFLTNTLLFKGFFTHGHWYNDLLV 69
DB 242 LEYKKQMTQRAASYSDDSYFGYSVAVGEFSEDAEDFVV-----GYPKNGTYGVYTI 295
QY 70 DLGSTAATSEYEGSSVDLYGQYGCAGTGNKTA---CMYGGVTLHD---NNRITTEK 123
DB 296 -LNGTDLRSLYNFSGQGM-ASYFGYSVSATDLNSDGLDLDLLICAPLFMDRTHGRVQEVG 353
QY 124 KVPINLWIDGKQTPIDKVKTSKKEVTVOELDLQARHYLHGKFG-----LYNSDS---- 174
DB 354 RVVYVLQGDHMESTPHL-ILTGMEE-----YGRFGSSIASLGLDLDQDGFN 397
QY 175 -----FGKVVORGLI-VFHSSEGSTVSVDLFDAGQ----- 204
DB 398 DIAIAPAPGGEAQRGAFFINQPGGVDSKPSQVLOGMGSSQSPFFGLSTRGHDLDG 457
QY 205 --YPTDLLRIYRDNNTI-----SSTSLSSIS 227
DB 458 NGYPDLIVGAFGVDTILVYRGRPIIHASASLSIS 491

RESULT 23
PRTA,ASPNG
ID PRTA,ASPNG STANDARD; PRT; 282 AA.
AC P24665;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Aspergillopepsin II precursor (EC 3.4.23.19) (Acid protease A)
DE (Protease A).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Var. Macrosporus;
RX MEDLINE=92011747; PubMed=1918060;
RA Inoue H., Kimura T., Makabe O., Takahashi K.;
RT "The gene and deduced protein sequences of the zymogen of Aspergillus
RT niger acid proteinase A.";
RL J. Biol. Chem. 266:19484-19489(1991).
RN [2]
RP SEQUENCE OF 60-98 AND 110-282.
RC STRAIN=Var. Macrosporus;
RX MEDLINE=92011746; PubMed=1918059;
RA Takahashi K., Inoue H., Sakai K., Kohama T., Kitahara S.,
RA Takishima K., Tanji M., Athauda S.B.P., Takahashi T., Akanuma H.,
RA Mamiya G., Yamasaki M.;
RT "The primary structure of Aspergillus niger acid proteinase A.";
RL J. Biol. Chem. 266:19480-19483(1991).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage in A chain of insulin:
CC 3-Asn-Gln-4, 13-Gly-I-Ala-14, and 26-Tyr-I-Thr-27.
CC -!- SUBUNIT: HETERODIMER OF TWO NONCOVALENTLY BOUND LIGHT AND HEAVY
CC CHAINS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A4.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)

```

```
CC EMBL; M68871; AAA32687.1; -.
DR PIR; A41025; A41025.
DR MEROPS; A04_002; -.
DR InterPro; IPR000250; AspartaseA4.
DR Pfam; PF01828; Peptidase_A4; 1.
DR PRINTS; PR00977; SCYTLDPASE.
DR ProDom; PD018627; SCYTLDPASE_A4; 1.
KW Hydrolase; Aspartyl protease; Zymogen; Signal.
FT SIGNAL 1 18
FT PROPEP 19 59
FT CHAIN 60 98
FT PROPEP 99 109
FT CHAIN 110 282
FT MOD_RES 110 110
FT DISULFID 115 139
FT DISULFID 127 210
SQ SEQUENCE 282 AA; 29887 MW; 4EA727F9AE33F72A CRC64;

Query Match 6.9%; Score 83.5; DB 1; Length 282;
Best Local Similarity 19.2%; Pred. No. 3.5;
Matches 52; Conservative 34; Mismatches 102; Indels 83; Gaps 10;

QY 13 RKSELOQTALGNLKOIYYNSKAITSEKSDAQFTNTLLFKGFTTGHWPWYNDLLVDLG 72
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 28 RKEARAAGKRHSNPPYIPGSDKEILKNGTNTNEEYSSN-----WAGAVLIGDG 75

QY 73 STAANTSEYEGSSVDL-----YGAYGY-----QCAG-----GTPNKTACMYGG 110
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 76 YTKVTGFTVPSVAGSSGSGGYGGYGNKRNKROSEECASAWVGIDGDTCTAILQRT 135

QY 111 VTLHDNRNLTKEKKVPINLWI-----DGKQTTVPIDKVKTSKKEVTVOE 154
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 136 VDF-----CYEGQTSYDANYEWPYDYDFSDITISEGDSIKVTVEATSKSGSATVEN 190

QY 155 LDL--QARHYLHGKF-----GLYNSDFGKGVORGLIVFHSSE-----GSTV 194
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 191 LTTGQSVTFTSFONVBGLCETNAEWIVDFESGSLVAFADFGSVTFTNABATSGGSTV 250

QY 195 SYDLFDAQOQYPTLLRIYRDNTTISSTLS 225
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 251 -----GPSDATVMDIEQDGSVLTFETSVS 273

RESULT 24
YB70_HAEIN STANDARD; PRT; 328 AA.
AC Q57527; O05047;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Hypothetical paba-like protein H1170.
GN H1170.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Keiley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
RL Science 269:496-512(1995).
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32797; AAC2834.1; -.
DR TIGR; H11170; -.
DR InterPro; IPR000350; Chorismate_bind.
DR Pfam; PF00425; chorismate_bind; 1.
DR ProDom; PD000779; Chorismate_bind; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 328 AA; 37734 MW; IC02BCD0088957B0 CRC64;

Query Match 6.8%; Score 82.5; DB 1; Length 328;
Best Local Similarity 27.3%; Pred. No. 5.2;
Matches 24; Conservative 16; Mismatches 35; Indels 13; Gaps 3;

QY 136 TTVPIDKVKTSKKEVT---VQELDLQARHYLHGKFLYNSDSFGKGVQRGLI-----V 185
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 237 TLLPAGISGAPKEKTTQIIQOAEKQKRGYTGIFGIDGKTLQSAVAIRFISQVDEKEY 296

QY 186 FHSSEGSTVSVDLFDAAQOQYPTLLRIY 213
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 297 FHSGGGITIH---SNAQDEYEELLEKVV 321

RESULT 25
YB70_BACSU STANDARD; PRT; 455 AA.
AC P25152;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical peptidase ywad precursor (EC 3.4.11.-).
GN YWAD OR IPA-8R.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95020537; PubMed=7934828;
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees.";
RL Mol. Microbiol. 10:371-384(1993).
RN [2]
RP SEQUENCE OF 1-68 FROM N.A.
RC STRAIN=168;
RX MEDLINE=92216127; PubMed=1806041;
RA Glaser P., Kunst F., Debarbouille M., Vertes A., Danchin A.,
RA Dedonder R.;
RT "A gene encoding a tyrosine tRNA synthetase is located near sacs in
RT Bacillus subtilis.";
RL DNA Seq. 1:251-261(1991).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL: X73124; CAA51564.1; -.
DR EMBL: X52480; CAA36725.1; -.
DR EMBL: Z99123; CAB15873.1; -.
DR PIR: S16427; S16427.
DR PIR: S39663; S39663.
DR MEROPS: M28; UPB; -.
DR Subtilisin; BG10554; ywad.
DR InterPro: IPR003137; PA.
DR Pfam: PF02225; PA; 1.
KW Hypothetical protein; Hydrolase; Amino peptidase; Metalloprotease;
KW Zinc; Signal; Complete proteome.
FT SIGNAL 1 31
FT CHAIN 32 455 HYPOTHETICAL PEPTIDASE YWAD.
SQ SEQUENCE 455 AA; 49450 MW; 89EE6A6EEB0CCE18 CRC64;

Query Match 6.8%; Score 82.5; DB 1; Length 455;
Best Local Similarity 22.0%; Pred. No. 7.9;
Matches 41; Conservative 25; Mismatches 67; Indels 53; Gaps 6;

QY 1 SEKSEINERDLRKSELOCTALGNLQIYYNSKAITSEKSADOF-----LTNTLLFK 55
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 53 SOLSEALGPR-----IAGTAEKKSALLIASSMRKLKLDVKVQRENIPDRLEGLTSSA 105
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 56 GFTTGHFWYNDLVLDGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACHYGGVTLHD 115
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 106 G-----RDILLQAASGAPTEEQGLTAPLYNAGLGYQ-----KDFADAKGKIALIS 152
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 116 NNRLT-----EKKVPINLMDGKQITVPIDKVK-----TSK 147
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 153 RGLTTYEKAKNAEAGAKAVIYNKRESLVPMPTNLNGKVGIPVYGKIKKDEALTOQ 212
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 148 KEVTQ 153
   | | | |
Db 213 KEATLK 218
```

```
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RN Science 277:1453-1474(1997).
RP [4]
RC SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshina T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN [5]
RP SEQUENCE OF 179-600 FROM N.A.
RX MEDLINE=94209210; PubMed=8157582;
RA Pruss B.M., Nels J.M., Park C., Wolfe A.J.;
RT "Mutations in NADH:ubiquinone oxidoreductase of Escherichia coli
RT affect growth on mixed amino acids.";
RL J. Bacteriol. 176:2143-2150(1994).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE. DOES COUPLE THE REDOX REACTION TO PROTON
CC TRANSLOCATION AND THUS CONSERVES THE REDOX ENERGY IN A PROTON
CC GRADIENT.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- SUBUNIT: COMPOSED OF 13 DIFFERENT SUBUNITS. SUBUNITS NUOCD, E,
CC F, AND, G CONSTITUTE THE PERIPHERAL SECTOR OF THE COMPLEX.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE COMPLEX I 30
CC KDA SUBUNIT FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE COMPLEX I 49
CC KDA SUBUNIT FAMILY.
CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 175.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68301; CAA48362.1; ALT_FRAME.
DR EMBL: X68301; CAA48363.1; ALT_FRAME.
DR EMBL: AE000317; AAC75346.1; -.
DR EMBL: D90859; BAA16115.1; -.
DR EMBL: D90860; BAA16120.1; ALT_INIT.
DR EMBL: L25055; AAA03535.1; ALT_INIT.
DR PIR: S37060; S37060.
DR PIR: S38312; S38312.
DR EcoGene: EG12084; nuoc.
DR InterPro: IPR001268; Complex1_30K.
DR InterPro: IPR001135; Oxidored_49kda.
DR Pfam: PF00329; complex1_30kd; 1.
DR Pfam: PF00346; complex1_49kd; 1.
DR ProDom: PD001581; Complex1_30K; 1.
DR PROSITE: PS00542; COMPLEX1_30K; 1.
DR PROSITE: PS00535; COMPLEX1_49K; 1.
KW Oxidoreductase; NAD; Ubiquinone; Multifunctional enzyme;
KW Complete proteome.
FT DOMAIN 1 190 NADH DEHYDROGENASE I CHAIN C.
FT DOMAIN 214 600 NADH DEHYDROGENASE I CHAIN D.
FT CONFLICT 370 370 H -> D (IN REF. 5).
FT CONFLICT 413 416 AYGA -> PWAR (IN REF. 1 AND 5).
FT CONFLICT 495 495 A -> R (IN REF. 5).
SQ SEQUENCE 600 AA; 68694 MW; 0092008D10243A19 CRC64;

Query Match 6.8%; Score 82.5; DB 1; Length 600;
```

Best Local Similarity 27.0%; Pred. No. 11;
Matches 51; Conservative 33; Mismatches 76; Indels 29; Gaps 12;

QY 26 LKQIYYNSKATSSSEKSDQFLTNLLPKGFTGHPWYN-DLLVDLGGTAATSEYEGSS 84
DB 380 LRFELDMFKRLASYEKAA---LONTLL-KGRSQGVAAAYGAKEALEWGTGAGLRATGID 435

QY 85 VDLVGA--YVGYO-----CAGGTEN-KTACMYGGVTLHDNNRLTPEE--KKVPINLW-I 131
DB 436 FDVRKARPYSGYENFDEIPVGGVSDCYTRVLMKVEELRQSLRILEOCLNNMPEGPFA 495

QY 132 DGKQTTVPIDKVKTSKEVTVQELDQARHYLHGKFG--LYNSDSFGG-KVORGLIVFH- 187
DB 496 DPLTTPP-----PKERTLOHETLTHFLQVSGVMPVMPANESQFMIEATKGINSYL 548

QY 188 SSEGSTVSY 196
DB 549 TSDGSTMYS 557

RESULT 27
DNAK_STRPY STANDARD; PRT; 607 AA.

AC P95831;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa protein) (HSP70).
GN DNAK OR SPY1760 OR SPYM18_1831.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 186103;
RN [1]
RP SEQUENCE FROM N.A.
RA Rioux C.R., Martin D., Hamel J., Brodeur B.R.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferreretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-WGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
RN [4]
RP SEQUENCE OF 1-11: 205-215 AND 426-436, AND MASS SPECTROMETRY.
RC STRAIN-JRS4 / Serotype M6;
RA Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
RA VanBogelen R.A.;
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes proteins.";
RN [5]
CC -!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC -!- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U72721; AAB39223.1; -;
DR EMBL; AE006604; AAK34501.1; -;
DR EMBL; AE010089; AAL98349.1; -;
DR HSSP; P04475; IDG4.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; Hsp70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation;
KW Complete proteome.
FT INIT_MET 0
FT MOD_RES 172 172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 607 AA; 64788 MW; 223BD4C737926F7F CRC64;

Query Match 6.7%; Score 82; DB 1; Length 607;
Best Local Similarity 27.9%; Pred. No. 12;
Matches 39; Conservative 20; Mismatches 57; Indels 24; Gaps 7;

QY 69 VDLGST-AATSEYEGSSVDLYGAYGYCAGGTPNKATCMYGGVTLHD---NNRLTPEKK 124
DB 6 IDLGTNSAVAVLEGTESKIIANPEGNRT---TPSVSEKNGEIIVGDAAKRQAVNPET 62

QY 125 VPINLWIDCKQTVPIDKVKTSKEVTVQELDQARHYLHGKFGLYNSDSFGCKVORGLI 184
DB 63 V---TSIKSKMGT--SEKVSANGKEYTPOEISAMILQYLKG-----YAEDYLGEKVEKAVI 113

QY 185 VFHSSEGSTVSVDLFDFAQQQ 204
DB 114 -----TVPAYFENDAQRQ 125

RESULT 28
PULA_THEMA
ID PULA_THEMA STANDARD; PRT; 843 AA.
AC O33840.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pullulanase precursor (EC 3.2.1.41) (Alpha-dextrin endo-1,6-alpha-glucosidase) (Pullulan 6-glucanohydrolase).
GN PULA OR TMI845.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OC Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MSB8 / DSM 3109;
RX MEDLINE=98115241; PubMed=9453151;
RA Bibel M., Brettl C., Gossler U., Kriegshaeuser G., Liebl W.;
RT "Isolation and analysis of genes for amylolytic enzymes of the hyperthermophilic bacterium Thermotoga maritima.";
RL FEMS Microbiol. Lett. 158:9-15(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,


```
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U72719; AAB39219.1; -
CC HSP; P04475; LDG4
CC InterPro; IPR001023; Hsp70.
CC Pfam; PF00012; HSP70; 1.
CC PRINTS; PR00301; HEATSHOCK70.
CC ProDom; PD000089; Hsp70; 1.
CC PROSITE; PS00297; HSP70.1; 1.
CC PROSITE; PS00329; HSP70.2; 1.
CC PROSITE; PS01036; HSP70.3; 1.
CC KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
CC MOD_RES 173 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC SEQUENCE 609 AA: 64980 MW; 298D8ADCC9D31E0C CRC64;
CC -----
CC Query Match 6.7%; Score 81; DB 1; Length 609;
CC Best Local Similarity 27.0%; Pred. No. 15;
CC Matches 37; Conservative 19; Mismatches 63; Indels 18; Gaps 5;
CC -----
CC QY 69 VDLGST-AATSEYEGSSVDLYGAYGYOCAGGTPNKACMGVTLHDNNRLTEKKVPI 127
CC DB 7 IDLGTNSAVALEGTESKIANPENGRNTPSPVSFKNGEIIIVGDAKQAVTNPDPT 63
CC -----
CC QY 128 NLWIDGKQTTVPIDKVTYSKKEVTVQELDLQARHYLHGKFLYNSDSFGKVGORGLIVFH 187
CC DB 64 VISIKSKMGT--SEKVSANGKEVTPQEISAMTLOVLKG----YAEDYLGEKVEKAVI--- 114
CC -----
CC QY 188 SSEGSTVSDLFDAQOQ 204
CC DB 115 -----TVPAYFNDARQ 126
CC -----
CC RESULT 31
CC WDRL_CAEEL
CC ID QRL176; STANDARD; PRT; 611 AA.
CC AC QRL176;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Actin interacting protein 1 (AIP1) (Uncoordinated protein 78).
CC GN UNC-78 OR C04F6.4.
CC OS Caenorhabditis elegans.
CC OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
CC OC Rhabditidae; Peloderinae; Caenorhabditis.
CC OX NCBI_TaxID=6239;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-Bristol N2;
CC RX MEDLINE=21157410; PubMed=11257131;
CC RA Ono S.;
CC RT "The Caenorhabditis elegans unc-78 gene encodes a homologue of
CC RT actin-interacting protein 1 required for organized assembly of muscle
CC RT actin filaments."
CC RL J. Cell Biol. 152:1313-1320(2001).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-Bristol N2;
CC RA Nhan M.;
CC RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Induces disassembly of actin filaments in conjunction
CC with ADF/cofilin family proteins. Regulator of actin organization
CC in myofibrils.
CC -!- SIMILARITY: CONTAINS 9 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE AIP1 FAMILY OF WD-REPEAT PROTEINS.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF324437; AAK11613.1; -
CC DR EMBL; U42835; AAA83588.1; -
CC DR WormPep; C04F6.4; C03924.
CC DR InterPro; IPR001680; WD40.
CC DR Pfam; PF00400; WD40; 9.
CC DR PRINTS; PR00320; GPROTEINBRPT.
CC DR ProDom; PD000018; WD40; 2.
CC DR SMART; SM00320; WD40; 10.
CC DR PROSITE; PS00678; WD_REPEATS_1; 3.
CC DR PROSITE; PS00082; WD_REPEATS_2; 6.
CC DR PROSITE; PS00294; WD_REPEATS_REGION; 1.
CC KW Actin-binding; Cytoskeleton; Repeat; WD repeat.
CC FT REPEAT 57 96
CC FT REPEAT 145 185
CC FT REPEAT 188 227
CC FT REPEAT 237 276
CC FT REPEAT 322 361
CC FT REPEAT 446 485
CC FT REPEAT 489 528
CC FT REPEAT 534 573
CC FT REPEAT 579 610
CC FT REPEAT 611 611
CC SQ SEQUENCE 611 AA: 65323 MW; 821452C661B5D27A CRC64;
CC -----
CC Query Match 6.7%; Score 81; DB 1; Length 611;
CC Best Local Similarity 18.7%; Pred. No. 15;
CC Matches 53; Conservative 46; Mismatches 78; Indels 106; Gaps 12;
CC -----
CC QY 23 LGNLKQIYYNKKATTSSEKSDAQELTNLLPKGFTH--PW-----YNDLLVDLGT 74
CC DB 313 LGSIDQVRYGHNKAITALSSAD---GKTLFSADAEGHINSWDISTGTSNRVFPDVHAT 368
CC -----
CC QY 75 AATSEYEGSSVDLYGAY-----GYOCAGGTPNK-----TA 105
CC DB 369 MITGKTKTSKGLFTVSWDDHLKVVYPAGSGVDSSKAVANKLSSQPLGLAVSADGDI 428
CC QY 106 CMYGGVTLHDNNRLTEKKVPI-----LWIDGKQTTVPIDKV---KTSKK 148
CC DB 429 ACYKHTAIYSHGKLT-----VPISYNSCVALSNDKQFVAVGDSKVHYVYKLSGASVSEV 485
CC QY 149 EYTVQELDL-----OAR-----HYLHGKFLYNSDSFGKVGORGLIVFHS 188
CC DB 486 KTVIHPAEITSVAFSNNNGAFLVATDQSRKVIPIYSVANNFELAHNTSW-----TPHT 536
CC QY 189 SEGSTVSDLFDAQOQYPTDLLRIYRDNPTTISSTLSISLYLY 231
CC DB 537 AKVACVSW-----PDNVRLATGSLDNLNVIV 563
CC -----
CC RESULT 32
CC PURA_METJA
CC ID PURA_METJA STANDARD; PRT; 345 AA.
CC AC Q57981;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)
CC DE (ADSS) (AMPSase).
CC GN PURA OR MJ0561.
CC OS Methanococcus jannaschii.
CC OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
CC OC Methanocaldococcaceae; Methanocaldococcus.
CC OX NCBI_TaxID=2190;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
```

```

RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geohagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RA jannaschii.";
RT Science 273:1058-1073(1996).
RL
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE DE NOVO PATHWAY OF PURINE
CC NUCLEOTIDE BIOSYNTHESIS.
CC -!- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
CC adenylosuccinate.
CC -!- PATHWAY: AMP biosynthesis; first committed step.
CC -!- SIMILARITY: BELONGS TO THE ADENYLOSUCCINATE SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U67505; AAB98554.1; -
CC HSP: P12283; IADE.
CC TIGR: MJ0361; -
CC InterPro: IPR001114; Asucc_synthase.
CC Pfam: PF00709; Adenylosucc_synt; 1.
CC ProDom: PD001188; Asucc_synthase; 1.
CC TIGRfam: TIGR00184; purA; 1.
CC PROSITE: PS01266; ADENYLOSUCCIN_SYN_1; 1.
CC PROSITE: PS00513; ADENYLOSUCCIN_SYN_2; FALSE_NEG.
KW Purine biosynthesis; Ligase; GTP-binding; Complete proteome.
FT NP_BIND 18 24 GTP (POTENTIAL).
FT ACT_SITE 144 144 BY SIMILARITY.
SQ SEQUENCE 345 AA; 37820 MW; 9974DBC30D1DEE72 CRC64;

Query Match 6.6%; Score 80; DB 1; Length 345;
Best Local Similarity 22.6%; Pred. No. 9.3;
Matches 44; Conservative 33; Mismatches 60; Indels 58; Gaps 11;

Qy 4 SEEINEK-DLRKKSLOQTALGNLQKQIYYNSKAITSSSEKADQFLTW----- 51
Db 166 SEEVNNAIDRGENVLIETQ-GTLLSLYYGYTPYVTSKDTTASSPAADVIGTPKVDVI 224

Qy 52 LIFKGFFT---GHPWYNLDLVLGSLTAATSEYEGSSVDLYG-----AYYGYOCAGG 99
Db 225 VVFKTFPTRVGAGPPPTMSLE-----EAESLGIVEYGTGTRRRRVGYDFELA-- 274

Qy 100 TPNKTACHYGGVT-----LHDNNRLTEKKVPINLWDGKQTTVPIDKV 143
Db 275 ---RKACRLNGATQIALGLDKYDCEYGVTEYNKLSKAKEFINK-IE-EVTGVPVTII 329

Qy 144 KTSKKEVTVQELDLQ 158
Db 330 STGPE--MHQTIDLR 342

RESULT 33
PCK_METH
ID PCK_METH STANDARD; PRT; 411 AA.
AC 027121;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PGK OR MTH1042.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;

```

```

OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN SEQUENCE FROM N.A.
RP SEQUENCE-DELTA H;
RC MEDLINE=98037514; PubMed=9371463;
RX Smith D.R., Doucette-Stamm L.A., Blakely D., Cook R., Gilbert K.,
RA Aldredge T., Bashirzadeh R., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Harrison D., Hoang L., Wang Y., Wierzbowski J., Gibson R.,
RA Spadafora R., Vicare R., Bush D., Safer H., Patwell D., Prabhakar S.,
RA Jiwan N., Caruso A., Goyal A., Pietrowski S., Church G.M.,
RA McDougall S., Shimer G., Rice P., Noelling J., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT deltaH; functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC -!- PATHWAY: Second phase of glycolysis; second step.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000876; AAB85538.1; -
CC HSP: P36204; 1VPE.
CC InterPro: IPR001576; PGK.
CC Pfam: PF00162; PGK; 1.
CC PRINTS: PR00477; PHGLYCKINASE.
CC PROSITE: PS00111; POLYGLYCERATE_KINASE; 1.
KW Transferase; Kinase; Glycolysis; Complete proteome.
FT ACT_SITE 411 AA; 44997 MW; E82D9737C6F74F76 CRC64;
SQ SEQUENCE 411 AA; 44997 MW; E82D9737C6F74F76 CRC64;

Query Match 6.6%; Score 80; DB 1; Length 411;
Best Local Similarity 22.5%; Pred. No. 11;
Matches 47; Conservative 35; Mismatches 75; Indels 52; Gaps 11;

Qy 7 INEKDLRKSSLOQTALGNLQKQIYYNSKAITSS-----EKSADQFLTWTLFK 55
Db 184 VMRELRL---TLQG-ALENRPVCYVILGVGVKDDSIMVMKNVLNGLSADLVLTGLVAN 239

Qy 56 GFFTGHPWYNLDLVLGSLTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGVTLLHD 115
Db 240 IFLAG-----CGVKIGKV-----NMDFIKS-RGY-----CDFIKVAKKL 272

Qy 116 NNRLTEKKVPINLW---DGKQTTVPIDKVTSK-KVTVQVELDQARH-----YHGG 165
Db 273 KKRFPERIVWPVDVAVCRDGRVDVPVKKIPNHPHQIDIGMETIKLYARRIREARTLFG 332

Qy 166 KFLYNSDSFGKVGQGLIVFHSSEGSTV 194
Db 333 PAGVFENPDFSIGTEDILNLAISSEGSFSI 361

RESULT 34
ITAL_HUMAN
ID ITAL_HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=9315124; PubMed=8428973;
RX Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit";
RL J. Biol. Chem. 268:2989-2996(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
CC E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 VFMA DOMAIN.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".
DR HSSP: P17301; IAOX.
DR Genew: HGNC:6134; ITGAL.
DR MIM: 192968;
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF00092; vwa; 1.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF01839; FG-GAP; 5.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; Vwf; 1.
DR PROSITE: PS0234; VWF_A; 1.
DR PROSITE: PS0234; VWF_A; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium.
FT DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMPM 114 1136 POTENTIAL.
FT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75 FG-GAP 1.
FT REPEAT 2 75 FG-GAP 2.
FT DOMAIN 147 360 VWF_A.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA_BIND 470 478 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GFFKR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;
Query Match 6.6%; Score 80; DB 1; Length 1151;
Best Local Similarity 27.3%; Pred. No. 41;
Matches 53; Conservative 25; Mismatches 66; Indels 50; Gaps 12;
Qy 72 GSTAATSEYEGSSV-----DLYGAYGYGOCAGTTPKNTACMYGGVTL----- 113
Db 596 GGDGKTLKFFGQSIHGMDLNG-----DGLTDVTTIGGLGAALFWSRDVAVYKVTMN 647
Qy 114 HDNNRLTEEEKVPINLWIDGKOT-----TVPIDKVKTSKKEVTVOELDLOARHYLHGKFG 168
Db 648 FEPKNVNIQKK---NCHWEGKETVCINATVCFE-VLKSKEDTIYEAOLQYVTLDSLRQ 703
Qy 169 LYNDSDFGG---KYQGLIVFHSSEGSTSVSYDLFDAQGGYPTDLLRIYRD-----NT 217
Db 704 ISRS-FFSGTQERKQVRNITV-RKSECTKHSFYMLD-KHDFODS-VRITLDNLNLTDPENG 759
Qy 218 TISSTSLISISLYLY 231
Db 760 PVLDDSLPNSVHEY 773
RESULT 35
YE64_PASMU
ID YE64_PASMU STANDARD; PRT; 324 AA.
AC Q9CKY6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical paba-like protein PMI464.
GN PMI464.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- SIMILARITY: TO THE C-TERMINAL OF PARA-AMINOBENZOATE SYNTHASE
COMPONENT I.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006183; AAK03548.1; -
CC InterPro; IPR000350; Chorisimate_bind.
CC Pfam; PF00425; chorisimate_bind; 1.
CC PRINTS; PR00095; ANTSNTHASEI.
CC PRODOM; PD000779; Chorisimate_bind; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 324 AA; 37309 MW; E9336FF3890A32AE CRC64;

```
Query Match          6.5%; Score 79.5; DB 1; Length 324;
Best Local Similarity 27.3%; Pred. No. 9.5;
Matches 24; Conservative 17; Mismatches 34; Indels 13; Gaps 3;

QY 136 TTVPIDKVKTSKKEVTQEL---DLQARHYLHGKGLYNSDFGKGVORGLI-----V 185
      | : | : | | | | : | : | | | | | : | : | : | : | : | : | : | : |
DB 236 TLLPAGSISGAPKKTGVIHAAEQRPGRYTGIFGLFDESLSQSAVAIRTEQVDERLI 295
      | : | : | | | | : | : | | | | | : | : | : | : | : | : | : | : |

QY 186 FHSSEGSTVSVLDLFAQGOYPTLLRIY 213
      | : | : | | | | : | : | | | | | : | : | : | : | : | : | : | : |
DB 296 FRSGGCTILSELED---EYELIQKV 320
      | : | : | | | | : | : | | | | | : | : | : | : | : | : | : | : |

RESULT 36
VEL_HPV35          STANDARD;          PRT;          637 AA.
ID VEL_HPV35
AC P27220:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Replication protein E1.
GN E1.
OS Human papillomavirus type 35.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10587;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92124753; PubMed=1310198;
RA Marich J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
RT "The phylogenetic relationship and complete nucleotide sequence of
human papillomavirus type 35.";
RL Virology 186:770-776(1992).
CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE. REQUIRED FOR INITIATION OF
CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74477; CRA52553.1; -.
DR EMBL; M74117; AAA46968.1; -.
DR PIR; A40824; W1WL35.
DR PIR; S36523; S36523.
DR InterPro; IPR001177; Papillom_E1.
DR Pfam; PF00519; El; 1.
DR Pfam; PF00524; ELN; 1.
KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW Nuclear protein.
FT NP_BIND 463 470 ATP (POTENTIAL).
FT DPVS -> SSV (IN REF. 2).
FT CONFLICT 31 34
FT CONFLICT 142 143 QQ -> HE (IN REF. 2).
FT CONFLICT 235 258 ESLTKIPKCYLIHIQCLCSWG ->
FT NFKHTIYVYINVRVHGA (IN REF. 2).
FT CONFLICT 269 281 CAKNRTTIEKLS -> VERREQQLTKIDA (IN REF.
FT 2).
FT CONFLICT 417 418 EK -> AQ (IN REF. 2).
FT CONFLICT 515 552 WAYIDQYLRNALDGNPISLDVKKALVOLKCPPLLINS ->
FT GIYPIEFKCTRWKYSISFRKALSIUHIIMPTFTYII (IN
FT REF. 2).
```


DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF00053; laminin_EGF_2.
DR Pfam; PF00054; laminin_G_2.
DR ProDom; PD002082; Laminin_G_2.
DR ProDom; PD003031; Laminin_B; 1.
DR ProSITE; PS00022; EGF_1; 19.
DR ProSITE; PS01186; EGF_2; 3.
DR ProSITE; PS01248; LAMININ_TYPE_EGF; 19.
DR ProSITE; PS50025; LAM_G_DOMAIN; 5.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin Egr-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 35
FT CHAIN 36 3695
FT DOMAIN 36 299 LAMININ ALPHA-5 CHAIN.
FT DOMAIN 300 358 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 359 428 LAMININ EGF-LIKE 1.
FT DOMAIN 429 474 LAMININ EGF-LIKE 2.
FT DOMAIN 494 540 LAMININ EGF-LIKE 3.
FT DOMAIN 541 586 LAMININ EGF-LIKE 4.
FT DOMAIN 587 631 LAMININ EGF-LIKE 5.
FT DOMAIN 632 676 LAMININ EGF-LIKE 6.
FT DOMAIN 677 722 LAMININ EGF-LIKE 7.
FT DOMAIN 723 775 LAMININ EGF-LIKE 8.
FT DOMAIN 776 828 LAMININ EGF-LIKE 9.
FT DOMAIN 829 850 LAMININ EGF-LIKE 10.
FT DOMAIN 851 1437 LAMININ EGF-LIKE 11 (INCOMPLETE).
FT DOMAIN 1438 1483 LAMININ DOMAIN IV 1 (DOMAIN IV B).
FT DOMAIN 1484 1527 LAMININ EGF-LIKE 12.
FT DOMAIN 1528 1576 LAMININ EGF-LIKE 13.
FT DOMAIN 1577 1627 LAMININ EGF-LIKE 14.
FT DOMAIN 1628 1637 LAMININ EGF-LIKE 15.
FT DOMAIN 1638 1830 LAMININ EGF-LIKE 16 (N-TERMINAL).
FT DOMAIN 1831 1863 LAMININ EGF-LIKE 17.
FT DOMAIN 1864 1912 LAMININ EGF-LIKE 18.
FT DOMAIN 1913 1968 LAMININ EGF-LIKE 19.
FT DOMAIN 1969 2022 LAMININ EGF-LIKE 20.
FT DOMAIN 2023 2069 LAMININ EGF-LIKE 21.
FT DOMAIN 2070 2116 LAMININ EGF-LIKE 22.
FT DOMAIN 2117 2166 LAMININ EGF-LIKE 23.
FT DOMAIN 2167 2335 LAMININ EGF-LIKE 24.
FT DOMAIN 2336 2929 LAMININ G-LIKE 1.
FT DOMAIN 2930 3115 LAMININ G-LIKE 2.
FT DOMAIN 3116 3292 LAMININ G-LIKE 3.
FT DOMAIN 3293 3340 LAMININ G-LIKE 4.
FT DOMAIN 3341 3513 LAMININ G-LIKE 5.
FT DOMAIN 3514 3692 LAMININ G-LIKE 6.
FT DOMAIN 3693 3921 LAMININ G-LIKE 7.
FT DOMAIN 3922 4221 LAMININ G-LIKE 8.
FT DOMAIN 4222 4400 LAMININ G-LIKE 9.
FT DOMAIN 4401 4577 LAMININ G-LIKE 10.
FT DOMAIN 4578 4711 LAMININ G-LIKE 11.
FT DOMAIN 4712 4944 LAMININ G-LIKE 12.
FT DOMAIN 4945 5155 LAMININ G-LIKE 13.
FT DOMAIN 5156 5266 LAMININ G-LIKE 14.
FT DOMAIN 5267 5388 LAMININ G-LIKE 15.
FT DOMAIN 5389 5500 LAMININ G-LIKE 16.
FT DOMAIN 5501 5621 LAMININ G-LIKE 17.
FT DOMAIN 5622 5711 LAMININ G-LIKE 18.
FT DOMAIN 5712 5844 LAMININ G-LIKE 19.
FT DOMAIN 5845 5999 LAMININ G-LIKE 20.

FT DISULFID 589 605 BY SIMILARITY.
FT DISULFID 607 616 BY SIMILARITY.
FT DISULFID 619 629 BY SIMILARITY.
FT DISULFID 632 644 BY SIMILARITY.
FT DISULFID 634 650 BY SIMILARITY.
FT DISULFID 652 661 BY SIMILARITY.
FT DISULFID 664 674 BY SIMILARITY.
FT DISULFID 677 689 BY SIMILARITY.
FT DISULFID 679 696 BY SIMILARITY.
FT DISULFID 698 707 BY SIMILARITY.
FT DISULFID 710 720 BY SIMILARITY.
Query Match 6.5%; Score 79; DB 1; Length 3695;
Best Local Similarity 29.4%; Pred. No. 2.1e-02;
Matches 58; Conservative 20; Mismatches 75; Indels 44; Gaps 13;
QY 56 GFTGHPWYNDLLVDLSTAAATSEYEGSSV-----DLYGAYGYQCAGGTNKT- 104
DB 41 GGFSLHPYPYFLAE--GARIAASATCGEAPARGSPRTEDLYCKLVGGPVAGGDPNQTI 98
QY 105 ACMYGGVTLHDNNRLTEKKVPINLWIDGKQ---TTVPIDK-VKTSKKEVTVOEDL-QA 159
DB 99 OGQYCDICTAANS-----NKAHPASNAIDGTERTWQSPPLSRGLEYNVNT---LDLQGV 151
QY 160 RH--YUHGKFG-----LYNSDSFGKVGRLVIFHSSEGSTVSYDLFDAGQGYPDT 208
DB 152 FHVAVYLIKFNANSPRDLWLERSMDFGRTYQWQF-FASSK-----RDLERFG--PQT 203
QY 209 LLRIYRDNNTTISSTLS 225
DB 204 LERITRDAACITTEYS 220
RESULT 41
YJRL_YEAST
ID YJRL_YEAST STANDARD; PRT; 396 AA.
AC P46992;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 43.0 kDa protein in CPSI-FPPI intergenic region.
GN YJL171C OR J0512.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OX Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rinke M., Domdey H.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YBR162C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 249446; CAAB9466.1; -.
KW SGD; S0003707; YJL171C.
KW Hypothetical protein.
SQ SEQUENCE 396 AA; 43014 MW; 279E858E7512670A CRC64;
Query Match 6.4%; Score 78.5; DB 1; Length 396;
Best Local Similarity 21.6%; Pred. No. 15;
Matches 64; Conservative 29; Mismatches 86; Indels 117; Gaps 16;
QY 26 LKQIYYNSKAIT-----SSEKSDAQF-----LNTTLF-K 55
DB 85 LKQAFYTPGFTVNNSSSDNNRLAYEYSSKTDADNTFLNHGGEASPCGLNALSAYS 144


```
FT CARBOHYD 632 632 O-LINKED (GLC...).
SQ SEQUENCE 762 AA; 82785 MW; 34EC9C784DECA67E CRC64;

Query Match
Best Local Similarity 6.4%; Score 78.5; DB 1; Length 762;
Matches 60; Conservative 28; Mismatches 105; Indels 49; Gaps 11;

QY 3 KSEINEKDLRKKSELOQTALGNLKOIYY-----NSKAITSEKSDAOFLLTLFLK 55
DB 90 KSEKSAFDPQNH-----WAVGQINLAYKLGIAQGVGNCKFDPNSELRYAQAALAFVLRAL 145
QY 56 GFETGHPWNLVDLSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKATACWYGGVTLHD 115
DB 146 GF-----KDLDPYG-----YLAKAQLD-GLVHGLNLA-----YNGVIKRG 180
QY 116 NNRLTSEK--KVPINLWIDGKOTTP--IDKVKTSKKEVTVOELDLQARHYLHGKFGLYN 171
DB 181 DLALILDLRALEVPWKYVDGKVEGLSEPLSKVAT--KAETVTIATNAQDSVEEGKVAVLD 239
QY 172 SDSFGKVGORGLVIFHSSEGSTVSYSYDLFDAQOQYPTDLLRIYRDNNTISST-----SLSIS 227
DB 240 KDGKLTINAGLVDFSEYLGKKV-----IVYSERFGDPVYVAEGDNDVVSFTSQDSVGT 295
QY 228 LY 229
DB 296 VY 297

RESULT 44
ADH1_ORYSA
ID ADH1_ORYSA STANDARD; PRT; 376 AA.
AC P20306;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Alcohol dehydrogenase 1 (EC 1.1.1.1).
GN ADH1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Indica-IR26; TISSUE=Seedling;
RX MEDLINE=93357433; PubMed=2562760;
RA Xie Y., Wu R.;
RT "Rice alcohol dehydrogenase genes: anaerobic induction, organ
RL Plant Mol. Biol. 13:53-68(1989).
CC -|- CATALYTIC ACTIVITY: An alcohol + NAD(+) -> an aldehyde or ketone +
CC NADH.
CC -|- COFACTOR: ZINC.
CC -|- SUBUNIT: HOMODIMER.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X16296; CAA34363.1; -
DR PIR; JQ0474; JQ0474.
DR HSSP; P11766; ITEX.
DR InterPro; IPR002328; ADH_zinc.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF00107; adh_zinc; 1.
DR PROSITE; PS00059; ADH_ZINC; 1.
KW Oxidoreductase; Zinc; NAD; Multigene family.
```

```
FT METAL 45 45 ZINC (CATALYTIC).
FT METAL 67 67 ZINC (CATALYTIC).
FT METAL 98 98 ZINC (SECOND ATOM).
FT METAL 100 100 ZINC (SECOND ATOM).
FT METAL 103 103 ZINC (SECOND ATOM).
FT METAL 111 111 ZINC (SECOND ATOM).
FT METAL 175 175 ZINC (CATALYTIC).
SQ SEQUENCE 376 AA; 40852 MW; B8ED6E025863D64F CRC64;

Query Match
Best Local Similarity 6.4%; Score 78; DB 1; Length 376;
Matches 47; Conservative 29; Mismatches 77; Indels 46; Gaps 7;

QY 3 KSEINEKDLRKKSELOQTALGNLKOIYYNSKAI-----TSSE-----KS 43
DB 104 KSAESNMCDLLINIDRGVMIGDGKSRFSINGKPIYHFGTSTFSEYTVMHVGVCAKINP 163
QY 44 ADOFTNTLLFGFTGHPWYNDLLVDLSTAAATSEYEGSSVDLYG-AYYGYQCAGGTPN 102
DB 164 AAPLDKVCVLSGGISTG-----LGATINVAKPKGSTVAIFGLGAVGLAAAGARI 213
QY 103 KTACWYGGVTLHDNRLTEKKVPINLWIDGKQTPVPIDKVKTSKKEVTVOELD----- 156
DB 214 RGASRIIGIDL--NANRFEEARKFGCTEFYNPKDHPKVOOVLA---EMTNGGVDRSVECT 269
QY 157 -----LQARHYLHGKFGFL 169
DB 270 GNINAMIOAFECVHDGWI 288

RESULT 45
DNAK_STRPN
ID DNAK_STRPN STANDARD; PRT; 607 AA.
AC P95829; O66035;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa
GN DNAK OR SP0517).
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RA Rioux C.R., Martin D., Hamel J., Brodeur B.R.;
RT "Heat shock protein HSP70 and amino terminus of DnaJ of Streptococcus
RT pneumoniae."
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Rx / CPl200;
RX MEDLINE=98231633; PubMed=9570114;
RA Kim S.-W., Choi I.-H., Kim S.-N., Kim Y.-H., Pyo S.-N., Rhee D.-K.;
RT "Molecular cloning, expression, and characterization of dnaK in
RT Streptococcus pneumoniae."
RL FEMS Microbiol. Lett. 161:217-224(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
```


RP SEQUENCE FROM N.A., AND SEQUENCE OF 34-41.
RC STRAIN-B6A-RI;
RX MEDLINE=94028940; PubMed=8215382;
RA Lee Y.-E., Lowe S.E., Zeikus J.G.;
RT "Gene cloning, sequencing, and biochemical characterization of
RT endoxylanase from *Thermoanaerobacterium saccharolyticum* B6A-RI.";
RL Appl. Environ. Microbiol. 59:3134-3137(1993).
RN [2]
RP ACTIVE SITE, AND MUTAGENESIS.
RX MEDLINE=93388520; PubMed=8376336;
RA Lee Y.-E., Lowe S.E., Henrissat B., Zeikus J.G.;
RT "Characterization of the active site and thermostability regions of
RT endoxylanase from *Thermoanaerobacterium saccharolyticum* B6A-RI.";
RL J. Bacteriol. 173:5890-5898(1993).
RN [2]
CC -!- FUNCTION: ENDO-ACTING ENZYME THAT RANDOMLY CLEAVES THE INTERNAL
CC XYLIDISIDIC LINKAGES OF THE XYLAN BACKBONE, YIELDING
CC XYLIDISIDIC LINKAGES OF THE XYLAN BACKBONE, WHICH ARE FURTHER
CC HYDROLYSED TO XLOSE MOLECULES BY BETA-XYLOSIDASE (EC 3.2.1.37).
CC THE TEMPERATURE AND PH OPTIMUM FOR THIS ENZYME ARE 70 DEGREES
CC CELSIUS AND 5.5, RESPECTIVELY. REQUIRES AT LEAST THREE XLOSE
CC RESIDUES FOR CATALYTIC ACTIVITY. DOES NOT HAVE ACTIVITY AGAINST
CC XYLULOSE.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC linkages in xylans.
CC -!- PATHWAY: Xylan degradation.
CC -!- INDUCTION: BY XYLAN AND XLOSE.
CC -!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
CC HYDROLASES).
CC -!- SIMILARITY: CONTAINS 2 S-LAYER HOMOMOLOGY (SLH) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M97882; AAA21812.1; ALT_SEQ.
DR HSP; P10478; 1XVZ.
DR InterPro; IPR003305; CBM_Cenc.
DR InterPro; IPR001547; GH_5.
DR InterPro; IPR001000; Glyco_hydro_10.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00331; Glyco_hydro_10; 1.
DR Pfam; PF00395; SLH; 2.
DR Pfam; PF02018; CBM_4_9; 2.
DR PRINTS; PR00134; GLYDRLASE10.
DR PROSITE; PS00391; GLYCOSYL_HYDROL_F10; 1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
DR PROSITE; PS01072; SLH_DOMAIN; 2.
DR Xylan degradation; Hydrolase; Glycosidase; Signal; Repeat.
KW SIGNAL 1 33
FT CHAIN 34 1157 ENDO-1,4-BETA-XYLANASE A.
FT DOMAIN 413 649 CATALYTIC (POTENTIAL).
FT DOMAIN 1055 1113 SLH 1.
FT DOMAIN 1114 1157 SLH 2.
FT ACT_SITE 495 495 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 537 537
FT ACT_SITE 600 600 NUCLEOPHILE.
FT MUTAGEN 537 537 D->N: LOSS OF ACTIVITY.
FT MUTAGEN 600 600 E->Q: LOSS OF ACTIVITY.
FT MUTAGEN 602 602 D->N: LOSS OF ACTIVITY.
SQ SEQUENCE 1157 AA; 128379 MW; 51FA6004497EC58B CRC64;
Query Match 6.4%; Score 78; DB 1; Length 1157;
Best Local Similarity 17.3%; Pred. No. 62;
Matches 38; Conservative 50; Mismatches 80; Indels 52; Gaps 11;
QY 23 LGNLKQIYYNSKATSSSEKSAQDLTLNLLFKGFTGHPWVNDLLVDLGSSTAATSEYE- 81
DB 160 IGTKLTLYW-----VESPDPTLEYIYDVVV-----TTQNP-----IQGNVIANETFFEN 205

QY 82 -----GSSVDLYGAYGYQCAGG-----TPNKTCACMYGGVTLHDNNRLTEKKKVPINLW 130
DB 206 GNTSGWIGTSSWRAVYGAHSGDYSLTTGRTA-NNNGPSYDLTGKIVPGQYVNDWF 264
QY 131 ---IDGKQTTVPIDKVKTSKKEVTQVELDLQARHY-----LHGKFGLYNSDSFGG 177
DB 265 VKFVNGNDT-----EQIKATVVKATSKDNTYIQVNDNFANVKNKGWTEIKGSFTLLPVAD----- 316
QY 178 KVQRGLIVFHSSEGSTVSYS--DLFDAQGGYQPDTLRLRIYRD 215
DB 317 --YSGISIVVESQNTPLEYIIDFVIGISNNQITQND 354
RESULT 48
GTF2_STRDO
ID GTF2_STRDO STANDARD; PRT; 1592 AA.
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715;
RX MEDLINE=91123227; PubMed=1704006;
RA Abo H., Matsumura T., Ohta H., Fukui K., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase).";
RL J. Bacteriol. 173:989-996(1991).
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -!- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl}(N+1).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: DENTAL CARIES.
CC -!- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -!- SIMILARITY: TO OTHER GLYCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90213; BAA14241.1; -.
DR PIR; A38175; A38175.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; GH_70.
DR Pfam; PF01473; CW_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1592 GLYCOSYLTRANSFERASE-I.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.

```
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE)
SQ SEQUENCE 1592 AA; 176167 MW; BC0A66D079351BCF CRC64;

Query Match 6.4%; Score 78; DB 1; Length 1592;
Best Local Similarity 25.4%; Pred. No. 93;
Matches 48; Conservative 21; Mismatches 84; Indels 36; Gaps 8;

QY 34 SKAITSSEKSAOFLNTLLFGFFTGHPWYNLLV-DLGLGTAATSEYEGS-SVDLYGAY 92
: | : | | | | : | | : | | | | | | | |
Db 28 ASALGASVASADTDASDDSNQTVVTGDTNNQATDOTSIAATATSEQASD--AATD 85
: | : | | | | : | | : | | | | | | | |

QY 93 GYOCAGGTGPNKACMYGGV-----TLIDNNRLTEKKVPINLWIDGKQTTVP 139
: | : | | | | | | | | | | | | | | | |
Db 86 QASAEQOTGTTASTDTAAQTNNANEAKWPTENOGFTDEMLAEAKNVAATESDIP 145
: | : | | | | | | | | | | | | | | | |

QY 140 IDKVKTSKKEVTQVQLDQARHYLHGKFLGNSDFGKQVQRGLIVFHSSEGSTVSYDLF 199
: | : | | | | | | | | | | | | | | | |
Db 146 SDLAKMS-----NVKQVD-----GKYYYDOD---GNVKNFAV---SVGDKIYY--F 185
: | : | | | | | | | | | | | | | | | |

QY 200 DAQGVQPD 208
: | | | | |
Db 186 DETGAYKDT 194
: | | | | |

RESULT 49
Y653_HAEIN
ID Y653_HAEIN STANDARD; PRT; 254 AA.
AC P44029;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative glycosyl transferase HI0653 (EC 2.-.-.-).
GN HI0653.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd."
RL Science 269:496-512(1995).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2. WARE/KDTX
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U32748; AAC22312.1; -.
CC TIGR: HI0653; -.
CC InterPro: IPR001173; Glycos_transf_2.
CC Pfam: PF00535; Glycos_transf_2; 1.
CC Hypothetical protein; Transferase; Glycosyltransferase;
KW
```

```
KW Complete proteome.
SQ SEQUENCE 254 AA; 29137 MW; 1F520A67EC07B149 CRC64;

Query Match 6.4%; Score 77.5; DB 1; Length 254;
Best Local Similarity 24.9%; Pred. No. 11;
Matches 48; Conservative 23; Mismatches 55; Indels 67; Gaps 12;

QY 42 KSADQFLNTLLFGFFTGHPWYNLLV-DLGSTAATSEYEGSSVDLYGA--YYGYQCAG 98
: | : | | | | | | | | | | | | | | | |
Db 11 KNEADLANCL-----DITVKDWVDEIIILDSGSTNTKEIALS-----YGAKFYENSQWG 61
: | : | | | | | | | | | | | | | | | |

QY 99 -GTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQTTVP-----IDKYVTKSKVEY 151
: | : | | | | | | | | | | | | | | | |
Db 62 FGKQRLAQOY-----VTSD-----YVLWLDADERTVTKLQQAIIILSAVK-NDRENT 106
: | : | | | | | | | | | | | | | | | |

QY 152 VOELDLQARHYLHGKFLGNSDFGKQVQRGLIVFHSSEGSTVSYDLFDAQGVQPDTLR 211
: | : | | | | | | | | | | | | | | | |
Db 107 VYEIPRV-----SEVFGREIR-----HS-----GMYPDYVVR 133
: | : | | | | | | | | | | | | | | | |

QY 212 IYRDNTTISSTSL 224
: | | | | | | | | | | | | | | | |
Db 134 LYRTWYQYNDSL 146
: | | | | | | | | | | | | | | | |

RESULT 50
MDR_PLAFF
ID MDR_PLAFF STANDARD; PRT; 1419 AA.
AC P13568;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein (Chloroquine resistance protein).
GN MDRI.
OS Plasmodium falciparum (isolate FC27 / Papua New Guinea).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5837;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89288297; PubMed=2701941;
RA Foote S.J., Thompson J.K., Cowman A.F., Kemp D.J.;
RT "Amplification of the multidrug resistance gene in some chloroquine-
RT resistant isolates of P. falciparum."
RL Cell 57:921-930(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92017800; PubMed=1922044;
RA Triglia T., Foote S.J., Kemp D.J., Cowman A.F.;
RT "Amplification of the multidrug resistance gene pfmdr1 in Plasmodium
RT falciparum has arisen as multiple independent events."
RL Mol. Cell. Biol. 11:5244-5250(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93149200; PubMed=8426608;
RA Wilson C.M., Volkman S.K., Thaithong S., Martin R.K., Kyle D.E.,
RA Milhous W.K., Wirth D.F.;
RT "Amplification of pfmdr1 associated with mefloquine and halofantrine
RT resistance in Plasmodium falciparum from Thailand."
RL Mol. Biochem. Parasitol. 57:151-160(1993).
CC -!- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- MISCELLANEOUS: P. FALCIPARUM RESISTANT TO THE DRUG CHLOROQUINE
CC HAVE MULTIPLE COPIES OF THE GENE CODING FOR MDR.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
```

```
DR EMBL; M29154; AAA29646.1; -
DR EMBL; X56851; CAA40180.1; -
DR EMBL; S53996; AAD13870.1; -
DR PIR; A32547; DVQOF.
DR PIR; S18204; S18204.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001140; ABCtransprtm.
DR Pfam; PF00005; ABC_tran; 2.
DR Pfam; PF00664; ABC_membrane; 2.
DR ProDom; PD000006; ABC_transportr; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.
FT DOMAIN 1 55 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 56 82 POTENTIAL.
FT TRANSMEM 91 116 POTENTIAL.
FT TRANSMEM 160 188 POTENTIAL.
FT TRANSMEM 194 212 POTENTIAL.
FT TRANSMEM 279 298 POTENTIAL.
FT TRANSMEM 314 338 POTENTIAL.
FT DOMAIN 339 788 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 789 807 POTENTIAL.
FT TRANSMEM 825 846 POTENTIAL.
FT TRANSMEM 908 928 POTENTIAL.
FT TRANSMEM 1028 1048 POTENTIAL.
FT TRANSMEM 1063 1083 POTENTIAL.
FT DOMAIN 1084 1419 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 643 661 POLY-ASN.
FT NP_BIND 413 420 ATP (POTENTIAL).
FT NP_BIND 1161 1168 ATP (POTENTIAL).
FT REPEAT 1 721
FT REPEAT 722 1419
FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1419 AA; 162251 MW; 0F96C7C1850B33D0 CRC64;

Query Match
Best Local Similarity 21.3%; Pred. No. 89;
Matches 47; Conservative 31; Mismatches 100; Indels 43; Gaps 8;

Qy 27 KOIYYN-----SKAITSSEKSADOFLTNTLLFKGFF---TGHWPYNDLLVDLGST 74
Db 1011 RAIDYKNGQKRRIIVNAALMGFSQSAQLFINSFAYWFGSFLIKRGTTILVDDFMKSLFTF 1070

Qy 75 AATSEYEGSSVDLYG-----AYGYQCA-----GGTPNKTACMYGGVTLHD 115
Db 1071 IPTGSYAGKLSLKGDSNAKLSFEKYYPLMIRKSNIDVRDDGGIRINKNLKFKGVDIKD 1130

Qy 116 -NNRLTEEEKVPI----NLWIDGKOTTVPIDKVKTSKK---EVTVQELDLQARHYL---- 163
Db 1131 VNFRIISRPNPVIYKNLSFTCDCKTTAIVGETSGKSTFMNLLRFDYDLKNDHIILKND 1190

Qy 164 HGKFGLYNSDSFGGKVGRLIVFHSSEGSTVSVDLFDAGQ 204
Db 1191 MTNFQDYQNNNNNSLVKKNVNEFSNQSGSAEDYTVFNNNGE 1231
```

Search completed: January 8, 2003, 11:57:17
Job time : 11.6343 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:54:52 ; Search time 15.4475 Seconds
(without alignments)
3107.877 Million cell updates/sec

Title: US-09-900-766-2
Perfect score: 1218
Sequence: 1 SEKSEFNEKDLRKSELQ.....RDNTTISLSISLYLVT 233

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	948	77.8	260	Q931M4	Q931M4 staphylococ
2	918	75.4	260	Q99SU3	Q99SU3 staphylococ
3	754	61.9	268	O85217	O85217 staphylococ
4	420.5	34.5	258	Q9EZM4	Q9EZM4 staphylococ
5	372.5	30.6	260	Q99T46	Q99T46 staphylococ
6	372.5	30.6	261	Q9EZM8	Q9EZM8 staphylococ
7	355	29.1	179	Q52075	Q52075 staphylococ
8	311	25.5	240	Q9F0L7	Q9F0L7 staphylococ
9	308	25.3	217	Q8RR76	Q8RR76 staphylococ
10	308	25.3	241	Q53585	Q53585 staphylococ
11	296.5	24.3	242	O85383	O85383 staphylococ
12	294.5	24.2	225	Q8VWV1	Q8VWV1 streptococ
13	293.5	24.1	225	Q91921	Q91921 streptococ
14	288.5	23.7	218	Q8RR75	Q8RR75 staphylococ
15	288.5	23.7	239	O05157	O05157 staphylococ
16	287.5	23.6	225	Q99Z21	Q99Z21 streptococ

17	287.5	23.6	239	2	Q06535	Q06535 staphylococ
18	284.5	23.4	239	2	Q06531	Q06531 staphylococ
19	284	23.3	239	2	Q9EZM7	Q9EZM7 staphylococ
20	282	23.2	239	16	Q99T47	Q99T47 staphylococ
21	278.5	22.9	256	2	O8VLM7	O8VLM7 staphylococ
22	277.5	22.8	239	2	Q06532	Q06532 staphylococ
23	277.5	22.8	239	2	Q06533	Q06533 staphylococ
24	276.5	22.7	234	2	Q9R5X4	Q9R5X4 staphylococ
25	276.5	22.7	242	2	Q93CC6	Q93CC6 staphylococ
26	265.5	21.8	271	2	Q9F0L6	Q9F0L6 staphylococ
27	264.5	21.7	239	2	Q53678	Q53678 staphylococ
28	263	21.6	242	2	O54476	O54476 staphylococ
29	263	21.6	242	2	Q93CC5	Q93CC5 staphylococ
30	258.5	21.2	239	2	Q06534	Q06534 staphylococ
31	254.5	20.9	236	2	Q97163	Q97163 streptococ
32	253.5	20.8	236	2	Q54779	Q54779 streptococ
33	252.5	20.7	222	2	Q9R931	Q9R931 streptococ
34	251.5	20.6	236	2	Q54696	Q54696 streptococ
35	251	20.6	260	2	Q54739	Q54739 streptococ
36	251	20.6	260	2	Q54971	Q54971 streptococ
37	251	20.6	260	2	O54738	O54738 streptococ
38	250	20.5	222	2	Q95524	Q95524 streptococ
39	250	20.5	222	2	Q938P4	Q938P4 streptococ
40	249	20.4	258	2	Q9ZNF2	Q9ZNF2 staphylococ
41	248.5	20.4	236	2	O57453	O57453 streptococ
42	241	19.8	233	2	Q8RR77	Q8RR77 staphylococ
43	241	19.8	258	2	Q9EZM3	Q9EZM3 staphylococ
44	241	19.8	258	16	O85382	O85382 staphylococ
45	238.5	19.6	259	2	Q936G4	Q936G4 staphylococ

ALIGNMENTS

RESULT 1

Q931M4
ID Q931M4 PRELIMINARY; PRT; 260 AA.
AC Q931M4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Enterotoxin P.
GN SEP OR SAV1948.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58110.1;
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin; 1.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 260 AA; 30016 MW; 15C2D36270FA8241 CRC64;

Query Match 77.8%; Score 948; DB 16; Length 260;
Best Local Similarity 76.4%; Pred. No. 4.3e-71;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

```
QY 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNSKAITSEKSADQFLNTLLFKGFTG 60
Db 28 SEKSEINEKDLRKSELQGTALGNLKOIYYNEAKTENKESHQDLQHTILFKGFTD 87
QY 61 HPWYNDLLVDLGSSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLT 120
Db 88 HSWYNDLLVDFDSKDTVDYKKGKVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLT 147
QY 121 EEKKVPINLWIDGKQTVTPIDVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVQ 180
Db 148 EEKKVPINLWLOGKQNTVPLETVKTKNKNVTVOELDLQARHYLHGKFGLYNSDSFGKVQ 207
QY 181 RGLIVFHSSEGSTVSYDLFDAQGOQPDTLRLIYRDNTTISSTLSLSLYTT 233
Db 208 RGLIVFHTSTEPSVNYDLFGAOGQYSNTLLRIYRDNTKINSENHMHIDIYLYTS 260

RESULT 2
Q99SU3 PRELIMINARY: PRT: 260 AA.
AC Q99SU3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Enterotoxin P.
SEP OR SA1761.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hoshino A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF003135; BAB43036.1; -.
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph_strep_toxin.
DR Pfam; PF01123; Staph_strep_toxin; 1.
DR Pfam; PF02876; Staph_strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Complete proteome.
SQ SEQUENCE 260 AA; 29708 MW; 087C584EC028CFDB CRC64;

Query Match 75.4%; Score 918; DB 16; Length 260;
Best Local Similarity 73.0%; Pred. No. 1.4e-68;
Matches 170; Conservative 21; Mismatches 42; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNSKAITSEKSADQFLNTLLFKGFTG 60
Db 28 SEKSEINEKDLRKSELQGTALSNLRQYHYNGSAIENKESNDQFLKNTLFDFFTG 87
QY 61 HPWYNDLLVDLGSSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLT 120
Db 88 HSWYNDLLVDSKDTANYIKGKKVDLYGYGYOCTGTPKTCMYGGVTLHNNQLE 147
QY 121 EEKKVPINLWIDGKQTVTPIDVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVQ 180
Db 148 EEKKVPINLWLOGKQNTVPLGTVKTKNKNVTVOELDLQARHYLHETYNLYNTDAENGKIQ 207
QY 181 RGLIVFHSSEGSTVSYDLFDAQGOQPDTLRLIYRDNTTISSTLSLSLYTT 233
Db 208 RGLIHFPPSSGDSVGYDLFGAOGQYPTQLRIYRDNTIKSKNMHIDIYLYTT 260
```

```
RESULT 3
O85217 PRELIMINARY: PRT: 268 AA.
AC O85217;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Enterotoxin J.
GN SEJ.
OS Staphylococcus aureus.
OG Plasmid pTB485.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KS11410;
RC MEDLINE=99052098; PubMed=9835033;
RA Zhang S., Iandolo J.J., Stewart G.C.;
RT "The enterotoxin D plasmid of Staphylococcus aureus encodes a second
RT enterotoxin determinant (sej).";
RL FEMS Microbiol. Lett. 168:227-233(1998).
DR EMBL; AF053140; AAC78590.1; -.
DR HSSP; P13163; 1ESF.
DR InterPro; IPR001961; Staph_strep_toxin.
DR Pfam; PF01123; Staph_strep_toxin; 1.
DR Pfam; PF02876; Staph_strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Plasmid.
SQ SEQUENCE 268 AA; 31230 MW; AC5F3546060ACE22 CRC64;

Query Match 61.9%; Score 754; DB 2; Length 268;
Best Local Similarity 59.3%; Pred. No. 6.3e-55;
Matches 137; Conservative 38; Mismatches 56; Indels 0; Gaps 0;

QY 3 KSEETNEKDLRKSELQGTALGNLKOIYYNSKAITSEKSADQFLNTLLFKGFTGHP 62
Db 27 KNETIKENLHKKSELSSITLNNLRHIYFFNEKISEKIMTEDQFLDYTLFKSFFISHS 86
QY 63 WYNDLLVDLGSSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLTEE 122
Db 87 QYNDLLVQFDSKETVYNKPKGQVDLYGSYGQCSCGKPKNTACMYGGVTLHNNQLYDT 146
QY 123 KKVPLNLWIDGKQTVTPIDVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVQRG 182
Db 147 KKIPLNLWIDSIRTVPVPLDIVTKNKKVTIOELDLQARYLHKQYNLXNPSTFDGKIQKG 206
QY 183 LIVFHSSEGSTVSYDLFDAQGOQPDTLRLIYRDNTTISSTLSLSLYTT 233
Db 207 LIVFHTSKEPLVSYDLFNVIGQYDPKLLKIYQDNKIIESENHMHIDIYLYTS 257

RESULT 4
Q9EZM4 PRELIMINARY: PRT: 258 AA.
ID Q9EZM4
AC Q9EZM4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Sek (Enterotoxin Sen).
GN SEK OR SEN OR SAVI825 OR SA1643.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC SPECIES=S.aureus; STRAIN=A900322;
RX MEDLINE=20571956; PubMed=11123352;
RA Jarraud S., Peyrat M.A., Lim A., Tristan A., Bes M., Mougel C.,
RA Etienne J., Vandenesch F., Bonneville M., Lina G.;
RT "egc, A highly prevalent operon of enterotoxin gene, forms a putative
RT nursery of superantigens in Staphylococcus aureus.";
RL J. Immunol. 166:669-677(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL: AF285760; AAG36956.1; -.
DR EMBL: AP003363; BAB57987.1; -.
DR EMBL: AP003135; BAB42911.1; -.
DR HSSP: P13163; 1EST.
DR InterPro: IPR001961; Staph_Strep_toxin.
DR Pfam: PF01123; Staph_Strep_toxin; 1.
DR PRINTS: PR00279; BACTRLTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
KW Complete proteome.
SQ SEQUENCE 258 AA; 29676 MW; 8A6C074F3E1F82D2 CRC64;
Query Match 34.5%; Score 420.5; DB 16; Length 258;
Best Local Similarity 38.0%; Pred. No. 3.5e-27;
Matches 87; Conservative 50; Mismatches 88; Indels 7; Gaps 4;
QY 6 EINEKDLRKSELOGTALGNLKOIYYNS-KAITSSEK-SADQFLNTLLFKGFFTHCHPWY 64
Db 32 EVDKKDLKKSLDSSKLFNLISYTDITWQLDENSKISTDQLLNTILKNIDISVLKT 91
QY 65 NDLVLDLSTAESEGSVDLYGAYGYCAGGTPNKACMYGGVTLHDNNRLTEKK 124
Db 92 SSLKVEFNDSLANQFKGNIDYGLYFGNKGCVGLTEKTSGLYGVTVTHDGNQLDDEKV 151
QY 125 VPTNLWIDCKQTVPTIDKVKTSKKEVTQELDLQARHYLHGKFLGYNLSDSFGKVGQRLI 184
Db 152 IGVNFKDGVQEGFV--IKTKAKVTQVELDKVRFKLENLYKIYNKDT--GNIQKGC 207
QY 185 VFHS--SEGTSVSYDLFDAQQGYPTLLRIYRDNTTISSTLSISLYLY 231
Db 208 FFHSHNHQDSFYDLYNNKSGVGAEEFQFYSNRTVSSSNVHIDVFLY 256
RESULT 5
Q99T46 PRELIMINARY; PRT; 260 AA.
AC Q99T46;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Enterotoxin SE0.
GN SE0 OR SAV1830 OR SA1648.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878; 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
```

```
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003363; BAB57992.1; -.
DR EMBL: AP003135; BAB42916.1; -.
DR HSSP: P13163; 1SXT.
DR InterPro: IPR001961; Staph_Strep_toxin.
DR Pfam: PF01123; Staph_Strep_toxin; 1.
DR PRINTS: PR00279; BACTRLTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
KW Complete proteome.
SQ SEQUENCE 260 AA; 29836 MW; AC927DDCC3E3B8EF CRC64;
Query Match 30.6%; Score 372.5; DB 16; Length 260;
Best Local Similarity 38.8%; Pred. No. 3.5e-23;
Matches 93; Conservative 32; Mismatches 92; Indels 23; Gaps 7;
QY 8 NEKD-----LRKKSELOGTALGNLKOIYYNS-KAITSSEK-SADQFLNTLLFKGFFTHGH 61
Db 29 NEEDPKIESLCKKSSVDPIALHINDDYNNRFTTVKSVSTEXFLDFDLFKSI---- 84
QY 62 PW-----YNDLLVDLSTAESEGSVDLYGAYGYCAGGTPNKACMYGGVTLH 114
Db 85 NWLDGTSAEFKDLKVEFSSAISKEFLGKTVDIYGVYKAKHGEHQVDVACTYGGVTPH 144
QY 115 DNNRLTEKKVPTNLWIDCKQTVPTIDKVKTSKKEVTQELDLQARHYLHGKFLGYNLSDS 174
Db 145 ENKUSEPKNICGVAYKDNVNVNTFI--VTDKKKVTQAELELDIKVKTKLNNAKLY--DR 200
QY 175 FGCKVQRGVLIVFHSSEGSTVS--YDLFDAQQGYPTLLRIYRDNTTISSTLSISLYLYT 232
Db 201 MTSDVQKGVYKFHSEHSEKSEFYDLYFKGNLPDQYQLIYNDNKTIDSSDIHIDVLYFT 260
RESULT 6
Q9EZM8 PRELIMINARY; PRT; 261 AA.
AC Q9EZM8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE SE0.
GN SE0.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A900322;
RX MEDLINE=20571956; PubMed=11123352;
RA Jarraud S., Peyrat M.A., Lim A., Tristan A., Bes M., Mougel C.,
RA Etienne J., Vandenesch F., Bonneville M., Lina G.;
RT "egc, A highly prevalent operon of enterotoxin gene, forms a putative
RT nursery of superantigens in Staphylococcus aureus.";
RL J. Immunol. 166:669-677(2001).
DR EMBL: AF285760; AAG36951.1; -.
DR HSSP: P13163; 1SXT.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam: PF01123; Staph_Strep_toxin; 1.
DR PRINTS: PR00279; BACTRLTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
```

SQ SEQUENCE 261 AA; 29949 MW; 24DA4DDD766288EC CRC64;

Query Match 30.6%; Score 372.5; DB 2; Length 261;
Best Local Similarity 38.8%; Pred. No. 3.5e-23;
Matches 93; Conservative 32; Mismatches 92; Indels 23; Gaps 7;

QY 8 NEKD-----LRKKSELQGTALGNLQIYYNS-KAITSEKSADQFLNTLLFKGFFTG 61
DB 30 NEEDPKIESLCKSSVDPAALHNDDYNNRFTTVKSVISTEXFLDPDLLFKSI---- 85

QY 62 PW-----YNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCACMGVTLH 114
DB 86 NWLDGISAEFKDLKVEFSSAISKEFLGTVDIYGVYKHAHGHEQVDTACTYGGVTPH 145

QY 115 DNNRLTEKKVPINLWIDGKQTTPIDKVKTSKKEVTQVELDQARHYLHGKFGLYNSDS 174
DB 146 ENNKUSEPNIGAVYKDNVWNTFI--VTTDKKVTQAEIDKIVKTRKLNNAKYLY--DR 201

QY 175 FGGQVQGLIVFHSSEGSTVS--YDLFDAQGYQPDTLRLIYRDNTTISTSLISLYLT 232
DB 202 MTSVDQKGYIKFHSSEHKESFYDLYFKGNLPDQYLOIYNDKNTIDSSDYHIDVYLET 261

RESULT 7
Q52075
ID Q52075 PRELIMINARY; PRT; 179 AA.

AC Q52075;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Enterotoxin D.
GN VIRC2.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89359112; PubMed=2549000;
RA Bayles K.W., Iandolo J.J.;
RT "Genetic and molecular analyses of the gene encoding staphylococcal enterotoxin D.";
RL J. Bacteriol. 171:4799-4806(1989).
DR EMBL; M94872; AAA98133.1; -.
DR HSSP: P13163; 1SXT.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam: PF01123; Staph_Strep_toxin; 1.
DR Pfam: PF02876; Staph_Strep_toxin; 1.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1;
KW Plasmid.
SQ SEQUENCE 179 AA; 20563 MW; 579FFB811BC08747 CRC64;

Query Match 29.1%; Score 355; DB 2; Length 179;
Best Local Similarity 45.3%; Pred. No. 6.1e-22;
Matches 68; Conservative 26; Mismatches 56; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELOGTALGNLQIYYNSKAITSEKSADQFLNTLLFKGFFTG 60
DB 26 NENIDSVREKELHKSEUSSTALNNKHYSADKNPTLIGENKSTGDOFLENTLLYKKFFTD 85

QY 61 HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCACMGVTLHDNNRLT 120
DB 86 LINFEDLLINFSNKEMAQHFKNVDVPIRYSINCYGGEIDRTACTYGGVTPHEGNKLIK 145

QY 121 EKKVPINLWIDGKQTTPIDKVKTSKKEV 150
DB 146 ERKKIPINLWINGQVEVSLDKVQTDKKNL 175

RESULT 8
Q9F0L7
ID Q9F0L7 PRELIMINARY; PRT; 240 AA.

Q9F0L7;
AC 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Sel (Extracellular enterotoxin L).
GN SEL OR SAV2008 OR SA1816.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX NCBI_TaxID=158878; 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=S. aureus; PubMed=111414901;
RX MEDLINE=20566668; PubMed=111414901;
RA Fitzgerald J.R., Monday S.R., Foster T.J., Bohach G.A., Hartigan P.J.,
RA Meaney W.J., Smyth C.J.;
RT "Characterization of a putative pathogenicity island from bovine Staphylococcus aureus encoding multiple superantigens.";
RL J. Bacteriol. 183:63-70(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Hiyama K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus.";
RL Lancet. 357:1225-1240(2001).
DR EMBL; AF217235; AAG29598.1; -.
DR EMBL; AP003364; BAB58170.1; -.
DR EMBL; AP003135; BAB43096.1; -.
DR HSSP: P13163; 1ESF.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam: PF01123; Staph_Strep_toxin; 1.
DR Pfam: PF02876; Staph_Strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Complete proteome.
SQ SEQUENCE 240 AA; 27496 MW; F40D62DA73197881 CRC64;

Query Match 25.5%; Score 311; DB 16; Length 240;
Best Local Similarity 33.3%; Pred. No. 4.1e-18;
Matches 74; Conservative 36; Mismatches 78; Indels 34; Gaps 6;

QY 24 GNLKQIY----YYNSKAITSEKSADQFLNTLLFKGFFTGHPWYND-LLVDLGSTAATS 78
DB 31 GNLRFYTKYEYVNLKVKDKNSPESHRL-----YSYKNDTLYAEFDNEYITS 79

QY 79 EYEGSSVDLYGAYGYQCAGGTPNKTCACMGVTLHDNNRLTEKKVPINLWIDGKQTIV 138
DB 80 DLKGNVDVFGISYK----GSNRT--YGGVTKAENKNKLDSPRIIPINLINGKHQTV 133

QY 139 PLDKVKTSKKEVTQVELDQARHYLHGKFGLY-----NSDSFGGKVGQGLIVF 186
DB 134 TTKSVSTDKMVTQAEIDVKLRKYLQDEENIYGHNDTCGKEYGTSKFYSGFDKGSVVF 193

QY 187 HSSEGSTVSVDLFDAGQGYPDTLRLIYRDNTTISTSLISL 228
DB 194 HMDGSGNSFYDLFTYGTGCLPESEFLKIYKDNKTVDSQFHLDV 235

RESULT 9
Q8RR76
ID Q8RR76 PRELIMINARY; PRT; 217 AA.
AC Q8RR76;

```

DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Enterotoxin H (Fragment).
GN SEH.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21871379; PubMed=11880405;
RA Omoie K., Ishikawa M., Shimoda Y., Hu D.L., Ueda S., Shinagawa K.;
RT "Detection of seg, seh, and sei genes in Staphylococcus aureus
RT isolates and determination of the Enterotoxin Productivities of S.
RT aureus Isolates Harboring seg, seh, or sei Genes.";
RL J. Clin. Microbiol. 40:857-862(2002).
DR EMBL; AB060536; BAB85990.1; -.
FT NON_TER
SQ SEQUENCE 217 AA; 25143 MW; A8A44E23F31DE80D CRC64;

Query Match 25.3%; Score 308; DB 2; Length 217;
Best Local Similarity 33.9%; Pred. No. 6.4e-18;
Matches 76; Conservative 46; Mismatches 90; Indels 12; Gaps 7;

QY 10 KDLRKKSELOGTALGNLKOIYYNSKATTSSEKSADQFLTNTLLFKGFTGHPWVNDLLV 69
DB 1 EDLHDKSELDTALAN--AYGOYNHPFKENIKSDEISGEKDLIFRN--QGDSG-NDLRV 55

QY 70 DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKACMYGGVTLHDNNRLTEKKVPINL 129
DB 56 KPATADLAQKFNKNVDIYGASFYKCEKISENISECLYGGTTL-NSEKLAQERVIGANV 114

QY 130 WIDGKQTTVPIDKVKTSKKEVTQVQLDQARHYLHGKFGLYNSDFSGKVQVORGLIVEHSS 189
DB 115 WVDGQKQETEL--ITNKKNVTQLQELDIKIRKILSDKYKIYKDS--EISKGLIEFDKM 169

QY 190 EGSTVSYDLFDAQQGYPTLLRIYRDNTTISSTLS-ISLYLYT 232
DB 170 TPRDYSFDIYDLKGENDYEDNKIYEDNKTLSDDISHIDVNLTY 213

RESULT 10
Q53585
ID Q53585 PRELIMINARY; PRT; 241 AA.
AC Q53585;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enterotoxin H precursor.
GN SEH.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=D4508;
RX MEDLINE=95053699; PubMed=7964453;
RA Ren K., Bannan J.D., Pancholi V., Cheung A.L., Robbins J.C.,
RA Fischetti V.A., Zabriskie J.B.;
RT "Characterization and biological properties of a new staphylococcal
RT exotoxin.";
RL J. Exp. Med. 180:1675-1683(1994).
DR EMBL; U11702; AAA19777.1; -.
DR HSP; P13163; IESF.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_tox_C; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal.
FT SIGNAL
FT CHAIN 1 24 POTENTIAL.
FT CHAIN 25 241 ENTEROTOXIN H.

```

```

SQ SEQUENCE 241 AA; 27858 MW; 70F77985877616CE CRC64;

Query Match 25.3%; Score 308; DB 2; Length 241;
Best Local Similarity 33.9%; Pred. No. 7.3e-18;
Matches 76; Conservative 46; Mismatches 90; Indels 12; Gaps 7;

QY 10 KDLRKKSELOGTALGNLKOIYYNSKATTSSEKSADQFLTNTLLFKGFTGHPWVNDLLV 69
DB 25 EDLHDKSELDTALAN--AYGOYNHPFKENIKSDEISGEKDLIFRN--QGDSG-NDLRV 79

QY 70 DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKACMYGGVTLHDNNRLTEKKVPINL 129
DB 80 KPATADLAQKFNKNVDIYGASFYKCEKISENISECLYGGTTL-NSEKLAQERVIGANV 138

QY 130 WIDGKQTTVPIDKVKTSKKEVTQVQLDQARHYLHGKFGLYNSDFSGKVQVORGLIVEHSS 189
DB 139 WVDGQKQETEL--ITNKKNVTQLQELDIKIRKILSDKYKIYKDS--EISKGLIEFDKM 193

QY 190 EGSTVSYDLFDAQQGYPTLLRIYRDNTTISSTLS-ISLYLYT 232
DB 194 TPRDYSFDIYDLKGENDYEDNKIYEDNKTLSDDISHIDVNLTY 237

RESULT 11
O85383
ID O85383 PRELIMINARY; PRT; 242 AA.
AC O85383;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Extracellular enterotoxin type I precursor (SEI).
GN SEI OR SAV1828 OR SA1646.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=S.aureus; STRAIN=FRI445;
RX MEDLINE=98298056; PubMed=9632603;
RA Munson S.H., Tremaine M.T., Betley M.J., Welch R.A.;
RT "Identification and characterization of staphylococcal enterotoxin
RT types G and I from Staphylococcus aureus.";
RL Infect. Immun. 66:3337-3348(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=S.aureus; STRAIN=A900322;
RX MEDLINE=20571956; PubMed=11123352;
RA Jarraud S., Peyrat M.A., Lim A., Tristan A., Bes M., Mougel C.,
RA Etienne J., Vandenesch F., Bonneville M., Lina G.;
RT "egc, a highly prevalent operon of enterotoxin gene, forms a putative
RT nursery of superantigens in Staphylococcus aureus.";
RL J. Immunol. 166:669-677(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF064774; AAC26661.1; -.
DR EMBL; AF285760; AAG36953.1; -.
DR EMBL; AP003363; BAB57990.1; -.
DR EMBL; AP003135; BAB42914.1; -.

```

```
DR HSP; P13163; LSXT.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_toxin; 1.
DR Pfam; PF02876; Strep_toxin; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal; Complete proteome.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 242 ENTEROTOXIN TYPE I.
SQ SEQUENCE 242 AA; 27863 MW; C5C8B4ACE5414A8 CRC64;

Query Match 24.3%; Score 296.5; DB 16; Length 242;
Best Local Similarity 32.4%; Pred. No. 6.7e-17;
Matches 77; Conservative 39; Mismatches 85; Indels 37; Gaps 10;

Qy 8 NEKDLRKKSELOG-TALGNLKOIY-----YNSKAITSEKSDQFLTWTLFKGFTGHP 62
Db 17 NIKDL---TYAGDITGVGNLRFNYTKHDYIDLKGVTDKNLP-----IANOLEFS---TG-- 64

Qy 63 WYNDLLVDLGSAAATSEYGVSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEE 122
Db 65 -TNDLISESNWDELSKPKGKGLDIFGIDYNGPC-----KSKYMYGGATL-SGOYLSA 116

Qy 123 KKVPTNLMDGQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFG-GK--- 178
Db 117 RKIPINLVNGKHKTISTDKIATNKKLVTAQIDVKLRRLQEEYNIYGHNTGKGEYV 176

Qy 179 -----VQRGLIVFHSSEGSTVSYDLFDAQGYPDTLRLRIYRDNTTISSTLSISL 228
Db 177 YSKSYGFENNGKVLFLHNNEKSFSDLYFTGDLGVPFLKIYEDNKTIIESEKFLDV 234

RESULT 12
Q8VWV1 PRELIMINARY; PRT; 225 AA.
AC Q8VWV1.
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Exotoxin I.
GN SPEI.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M1;
RX MEDLINE=21259899; PubMed=11359827;
RA Proft T., Arcus V.L., Handley V., Baker E.N., Fraser J.D.;
RT "Immunological and biochemical characterization of streptococcal
RT pyrogenic exotoxins I and J (SPE-I and SPE-J) from Streptococcus
RT pyogenes";
RL J. Immunol. 166:6711-6719(2001).
DR EMBL; AF438524; AAL31571.1; -.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_toxin; 1.
DR Pfam; PF02876; Strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; UNKNOWN_1.
SQ SEQUENCE 225 AA; 25654 MW; DB193667890ACDAF CRC64;

Query Match 24.2%; Score 294.5; DB 2; Length 225;
Best Local Similarity 32.4%; Pred. No. 8.9e-17;
Matches 72; Conservative 40; Mismatches 85; Indels 25; Gaps 8;

Qy 25 NLKOIY-YNSKAITSEKSDQFLTNTLLFKGFTGHPWYN---DLLVDLGSAAATSEY 80
Db 8 NLRLSYDPTVEVKINEGPPF-SGSLFYKNI-----PYGNSSTELKVELNSVEKANFF 62

Qy 81 EGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDG---KQTT 137
Db 63 SGRKVDIFTLEYSPPCNSNIKKNS---YGGITLSDGNRI-DKKNIPVNFIDGVQOKYSY 118

Qy 138 VPIDKVKTSKEVTVOELDLQARHYLHGKFLGY-----NSDSFGKVGQRLIVFHS 188
Db 119 TDISTVSTDKKEVTIQELDVKSRYYLQKHFNIYGVYKDFGRSSRFQSGFEGNIIFHL 178

Qy 189 SEGSTVSYDLFDAQGYPDTLRLRIYRDNTTISSTLSISL 230
Db 179 NSGERISYNLFDTHGHDRESMLKKYSDNKTAYSQDLHIDIYL 220

RESULT 14
Q8RR75 PRELIMINARY; PRT; 218 AA.
AC Q8RR75.
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Enterotoxin I (Fragment).
GN SEI.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
```

```
Qy 138 VPIDKVKTSKEVTVOELDLQARHYLHGKFLGY-----NSDSFGKVGQRLIVFHS 188
Db 119 TDISTVSTDKKEVTIQELDVKSRYYLQKHFNIYGVYKDFGRSSRFQSGFEGNIIFHL 178

Qy 189 SEGSTVSYDLFDAQGYPDTLRLRIYRDNTTISSTLSISL 230
Db 179 NSGERISYNLFDTHGHDRESMLKKYSDNKTAYSQDLHIDIYL 220

RESULT 13
Q9L921 PRELIMINARY; PRT; 225 AA.
AC Q9L921.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 25.7 kDa protein.
OS Streptococcus equi.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CF32;
RA Artushin S.C., Timoney J.F., Sheoran A.S.;
RT "Identification and molecular characterization of mitogens from
RT Streptococcus equi";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF186180; AAF72808.1; -.
DR HSP; P13163; LSXT.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_toxin; 1.
DR Pfam; PF02876; Strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
SQ SEQUENCE 225 AA; 25684 MW; EB43D47B26FB9F89 CRC64;

Query Match 24.1%; Score 293.5; DB 2; Length 225;
Best Local Similarity 32.4%; Pred. No. 1.1e-16;
Matches 72; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

Qy 25 NLKOIY-YNSKAITSEKSDQFLTNTLLFKGFTGHPWYN---DLLVDLGSAAATSEY 80
Db 8 NLRLSYDPTVEVKINEGPPF-SGSLFYKNI-----PYGNSSTELKVELNSVEKANFF 62

Qy 81 EGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDG---KQTT 137
Db 63 SGRKVDIFTLEYSPPCNSNIKKNS---YGGITLSDGNRI-DKKNIPVNFIDGVQOKYSY 118

Qy 138 VPIDKVKTSKEVTVOELDLQARHYLHGKFLGY-----NSDSFGKVGQRLIVFHS 188
Db 119 TDISTVSTDKKEVTIQELDVKSRYYLQKHFNIYGVYKDFGRSSRFQSGFEGNIIFHL 178

Qy 189 SEGSTVSYDLFDAQGYPDTLRLRIYRDNTTISSTLSISL 230
Db 179 NSGERISYNLFDTHGHDRESMLKKYSDNKTAYSQDLHIDIYL 220

RESULT 14
Q8RR75 PRELIMINARY; PRT; 218 AA.
AC Q8RR75.
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Enterotoxin I (Fragment).
GN SEI.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
```

```

RN  SEQUENCE FROM N.A.
RP  MEDLINE=21871379; PubMed=11880405;
RA  Omoe K., Ishikawa M., Shimoda Y., Hu D.L., Ueda S., Shinagawa K.;
RT  "Detection of seg, seh, and sei genes in Staphylococcus aureus
RT  Isolates and Determination of the Enterotoxin Productivities of S.
RT  aureus Isolates Harboring seg, seh, or sei Genes.";
RL  J. Clin. Microbiol. 40:857-862(2002).
DR  EMBL; AB060537; BAB85991.1; -.
FT  NON_TER
SQ  SEQUENCE 218 AA; 24909 MW; 101728FB71BF662 CRC64;

Query Match      23.7%; Score 288.5; DB 2; Length 218;
Best Local Similarity 31.5%; Pred. No. 2.7e-16;
Matches 70; Conservative 39; Mismatches 80; Indels 33; Gaps 8;

QY  23 LGNLKQIY----YVNSKAITSSSEKSDAQFVLTNTLLFKGFFTGHPWYNDLVLVGLGTAATS 78
DB  6 VGNLRFYTKHYIDUKGVTDKNLP-----IANQLEFS---TG---TNDLISENNWDEIS 55
QY  79 EYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNRNLTTEEKVPINLWIDGKQTV 138
DB  56 KFKGKKLDIFGIDYNGPC-----KSKYFEGGATL-SGOYLSARKIPINLWYNGKHKT 108
QY  139 PIDKVKTSKEVTVOELDQARHYLHGKFLYNSDSFG-GK-----VORGLIVF 186
DB  109 STDKIATNKKLTAQIBIDVKRLRYLQEEYNIYGHNNVTGKGYKSKFYSGFNNGKVL 168
QY  187 HSSEGSTVSVDLFDAGQVPTDLLRIYRDNNTTISSTLSISL 228
DB  169 HLNNEKSFYDLFTYDGDLPVSLFKIYEDNKIESEKFKHLDV 210

RESULT 15
O05157 PRELIMINARY; PRT; 239 AA.
AC  O05157;
DT  01-JUL-1997 (TrEMBLrel. 04, Created)
DT  01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE  Type C enterotoxin (Fragment).
OS  Staphylococcus intermedius.
OC  Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX  Staphylococcus.
ON  NCBI_TaxID=1285;
RN  [1]
RC  SEQUENCE FROM N.A.
RP  STRAIN=95-011195;
RA  Edwards V.M., Deringer J.R., Callantine S.D., Deobald C.F.,
RA  Berger P.H., Kapur V., Stauffacher C.V., Bohach G.A.;
RT  "Characterization of the type C enterotoxin (SEC-canine) produced by
RT  Staphylococcus intermedius pyoderma isolates.";
RL  Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U91526; AAB50248.1; -.
DR  HSSP; P23313; IJCK.
DR  InterPro; IPR001961; Stap/Strep_toxin.
DR  Pfam; PF01123; Stap_Strp_toxin; 1.
DR  Pfam; PF02876; Stap_Strp_tox.C; 1.
DR  PRINTS; PR00279; BACTRLTOXIN.
DR  PROSITE; PS00277; BACTRLTOXIN.
DR  PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR  PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT  NON_TER
SQ  SEQUENCE 239 AA; 27536 MW; D660644660DE4191 CRC64;

Query Match      23.7%; Score 288.5; DB 2; Length 239;
Best Local Similarity 34.5%; Pred. No. 3e-16;
Matches 81; Conservative 37; Mismatches 96; Indels 21; Gaps 9;

QY  11 DLKKKSELQATGALNKLQIY--YVNSKAITSSSEKSDAQFVLTNTLLFKGFFTGHPWYNDLL 68
DB  10 DLHKSEFTGT-MGNMKCLYDDYYS---ATKVKSDYKFLAHDLIYNSDKLKNYDKVK 65
QY  69 VDLGTAATSEYEGSSVDLYGAYGYOC-----AGGTPNKTACMYGGVTLHDNRNLT 121

```

```

DB  66 TELLNEDLAKYKDEVVDVYGSNYVNCYFSSKSDNVGKVTGGTKMYGTTKHEGNHFN 125
QY  122 E--KKVPINLWIDGKOTVPIDKVKTSSKEVTVOELDQARHYLHGKFLYNSDSFGKV 179
DB  126 GNLQNLIRVY-ENKRTNISFD-VQTDKKSVAQELDIKARNFLLINKNKLYEFNS--SPY 181
QY  180 QRGLIVFHSSEGSTVSVDLFDAGQVPTD--TLRLIYRDNNTTISSTLSISLYLT 232
DB  182 ETGYIKFIESNDNTFWYDMMPAPGDKFDQSKYLMYSDNKTVDKSKVKIEVHLTT 236

RESULT 16
Q99221 PRELIMINARY; PRT; 225 AA.
AC  Q99221;
DT  01-JUN-2001 (TrEMBLrel. 17, Created)
DT  01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE  Streptococcus exotoxin I.
DE  Streptococcus exotoxin I.
GN  SPEI OR SPY1007.
OS  Streptococcus pyogenes.
OC  Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OX  Streptococcaceae; Streptococcus.
ON  NCBI_TaxID=1314;
RN  [1]
RC  SEQUENCE FROM N.A.
RP  STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX  MEDLINE=21192684; PubMed=11296296;
RA  Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA  Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA  Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA  Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT  "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL  Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR  EMBL; AE006546; AAK3906.1; -.
DR  HSSP; P13163; ISXT.
DR  InterPro; IPR001961; Stap/Strep_toxin.
DR  Pfam; PF01123; Stap_Strp_toxin; 1.
DR  Pfam; PF02876; Stap_Strp_tox.C; 1.
DR  PRINTS; PR00279; BACTRLTOXIN.
DR  PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR  PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW  Complete proteome.
SQ  SEQUENCE 225 AA; 25611 MW; CE0D6736CC11CG04 CRC64;

Query Match      23.6%; Score 287.5; DB 16; Length 225;
Best Local Similarity 32.0%; Pred. No. 3.4e-16;
Matches 71; Conservative 40; Mismatches 86; Indels 25; Gaps 8;

QY  25 NLKQIY-YVNSKAITSSSEKSDAQFVLTNTLLFKGFFTGHPWYND--DLLVDLGSTAATSEY 80
DB  8 NLRLNSTYDPTVEVKGINEGPPF-SGSLFYKNI----PYGNSSIELKVELNSVEKANFF 62
QY  81 EGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNRNLTTEEKVPINLWIDG---KQTF 137
DB  63 SGKRVDIFLTLEYSPPCNSNIKNS---YGGITLSDGNRI-DKKNPINWIFIDGVOQKYSY 118
QY  138 VPIDKVKTSKEVTVOELDQARHYLHGKFLY-----NSDSFGKVGQGLIVFHS 188
DB  119 TDISTGSTDKKEVTIQELDVKSRYLQKHFNHYGFGVDKDFGRSSRFQSGFEENIIFHL 178
QY  189 SEGSTVSVDLFDAGQVPTDLLRIYRDNNTTISSTLSISLYL 230
DB  179 NSGERISYNFLDTGCHGDRSMLKKYSNDKNTAYSDQLHDIYL 220

RESULT 17
Q06535 PRELIMINARY; PRT; 239 AA.
ID  Q06535;
AC  Q06535;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

```

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Enterotoxin type C-3 (SEC3) (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX Staphylococcus.
RN NCBI_TaxID=1280;
RP [1]
RC SEQUENCE FROM N.A.
RX STRAIN=FRI 909;
RX MEDLINE=94011313; PubMed=8406814;
RA Marr J.C., Lyon J.D., Roberson J.R., Luper M., Bohach G.A.;
RT "Characterization of novel type C staphylococcal enterotoxins:
RT biological and evolutionary implications.";
RL Infect. Immun. 61:4254-4262(1993).
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC -!- STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- SIMILARITY: THE DIFFERENT S. AUREUS ENTEROTOXINS AND S. PYOGENES
CC PYROGENIC EXOTOXINS ARE ALL RELATED.
CC EMBL; L13377; AAA26621.1; -.
DR HSP; P23313; LJCK.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strp_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Superantigen.
FT NON_TER 1
SQ SEQUENCE 239 AA; 27648 MW; CC3CB3B04E4119E0 CRC64;

Query Match 23.6%; Score 287.5; DB 2; Length 239;
Best Local Similarity 33.3%; Pred. No. 3.7e-16;
Matches 78; Conservative 43; Mismatches 94; Indels 19; Gaps 9;

QY 11 DLRRKSELQCTALGNLKKQIYYNNSKAITSE--KSADQFLTNTLLFKGFTTGHWPYNDLLV 69
DB 10 DLHKSSEFTGT-MGNMK--YLYDDHYVSATKVKSVKFLAHLDIYINHKLNLYDKVKT 66
QY 70 DLGTAATSEYEGSSVDLYCAYGYQC-----AGTPTNKTCMYGGVTLHDNNRLTEE 122
DB 67 ELLNEDLAKNKYDEVVDVYGSNYYVNCYFSSKDNVCKVTSGKTCMYGGITKHEGHNFDNG 126
QY 123 --KKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYNSDFGKVKQ 180
DB 127 NLQNVLIRVY-ENKNTISFE-VQTDKSKSVTAQELDIKARFLINKKNLYEFSN--SPYE 182
QY 181 RGLVPHSSEGSTSVSYDLFDAQGOYPD--TLRIYRDNTTISSTLSISLYLYT 232
DB 183 TGYIKFIESNGNTFWYDMMPAPGDKFDQSKYLMYKDNKMWDSKVKIEVHLTT 236

RESULT 18
Q06531 ID Q06531 PRELIMINARY; PRT; 239 AA.
AC Q06531;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Enterotoxin type C-4 (SEC4446) (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX Staphylococcus.
RN NCBI_TaxID=1280;
RP [1]
RC SEQUENCE FROM N.A.
RX STRAIN=4446;
RX MEDLINE=94011313; PubMed=8406814;
RA Marr J.C., Lyon J.D., Roberson J.R., Luper M., Bohach G.A.;
RT "Characterization of novel type C staphylococcal enterotoxins:
RT biological and evolutionary implications.";
RL Infect. Immun. 61:4254-4262(1993).
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC -!- STAPHYLOCOCCAL FOOD POISONING SYNDROME.

```

```

CC -!- SIMILARITY: THE DIFFERENT S. AUREUS ENTEROTOXINS AND S. PYOGENES
CC PYROGENIC EXOTOXINS ARE ALL RELATED.
CC EMBL; L13374; AAA26618.1; -.
DR HSP; P34071; ISE2.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strp_toxin; 1.
DR Pfam; PF02876; Staph_Strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Superantigen.
FT NON_TER 1
SQ SEQUENCE 239 AA; 27612 MW; BB7BD6204731ED24 CRC64;

Query Match 23.4%; Score 284.5; DB 2; Length 239;
Best Local Similarity 32.5%; Pred. No. 6.5e-16;
Matches 76; Conservative 45; Mismatches 94; Indels 19; Gaps 9;

QY 11 DLRRKSELQCTALGNLKKQIYYNNSKAITSE--KSADQFLTNTLLFKGFTTGHWPYNDLLV 69
DB 10 DLHKSSEFTGT-MGNMK--YLYDDHYVSATKVKSVKFLAHLDIYINSDKRLKNYDKVKT 66
QY 70 DLGTAATSEYEGSSVDLYCAYGYQC-----AGTPTNKTCMYGGVTLHDNNRLTEE 122
DB 67 ELLNEDLAKNKYDEVVDVYGSNYYVNCYFSSKDNVCKVTSGKTCMYGGITKHEGHNFDNG 126
QY 123 --KKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYNSDFGKVKQ 180
DB 127 NLQNVLIRVY-ENKNTISFE-VQTDKSKSVTAQELDIKARFLINKKNLYEFSN--SPYE 182
QY 181 RGLVPHSSEGSTSVSYDLFDAQGOYPD--TLRIYRDNTTISSTLSISLYLYT 232
DB 183 TGYIKFIESNGNTFWYDMMPAPGDKFDQSKYLMYKDNKMTVDKSKVKIEVHLTT 236

RESULT 19
Q062M7 ID Q062M7 PRELIMINARY; PRT; 239 AA.
AC Q062M7;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE SEM.
GN SEM.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX Staphylococcus.
RN NCBI_TaxID=1280;
RP [1]
RC SEQUENCE FROM N.A.
RX STRAIN=A900322;
RX MEDLINE=20571956; PubMed=11123352;
RA Jarraud S., Peyrat M.A., Lim A., Tristan A., Bes M., Mougel C.,
RA Etienne J., Vandenesch F., Bonneville M., Lina G.;
RT "egc, A highly prevalent operon of enterotoxin gene, forms a putative
RT nursery of superantigens in Staphylococcus aureus.";
RL J. Immunol. 166:669-677(2001).
DR EMBL; AF285760; AAG36952.1; -.
DR HSP; P13163; LSXT.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strp_toxin; 1.
DR Pfam; PF02876; Staph_Strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
SQ SEQUENCE 239 AA; 27371 MW; 69240BE23C44028A CRC64;

Query Match 23.3%; Score 284; DB 2; Length 239;
Best Local Similarity 30.4%; Pred. No. 7.2e-16;
Matches 69; Conservative 38; Mismatches 80; Indels 40; Gaps 6;

QY 23 LGNLKKQIYYNNSKAIT--SSEKADQFLTNTLLFKGFTTGHWPYNDLLVDLGSTAA---- 76
DB 24 VGVNLNRNYGYSPIEDHQSNPENNHLSQLVFS-----MDNSTVTAEFN 70

```

RESULT 21					
ID	Q8VLW7	PRELIMINARY;	PRT;	256 AA.	
AC	Q8VLW7;				
DT	01-MAR-2002	(TrEMBLrel. 20, Created)			
DT	01-MAR-2002	(TrEMBLrel. 20, Last sequence update)			
DT	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)			
DE	Staphylococcal enterotoxin Q.				
GN	ENTQ.				
OS	Staphylococcus aureus.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;				
OC	Staphylococcus.				
OX	NCBI_TaxID=1280;				
RN	[1]				
RC	SEQUENCE FROM N.A.				
RC	STRAIN=RN4282;				
RX	MEDLINE=98385824; PubMed=9720870;				
RA	Lindsay J.A., Ruzin A., Ross H.F., Kurepina N., Novick R.P.;				
RA	"The gene for toxic shock toxin is carried by a family of mobile				
RT	pathogenicity islands in staphylococcus aureus."				
RL	Mol. Microbiol. 29:527-543(1998).				
RL	[2]				
RC	SEQUENCE FROM N.A.				
RC	STRAIN=RN4282;				
RA	Lindsay J.A., Kreiswirth B.N., Novick R.P.;				
RL	Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.				
RL	[3]				
RC	SEQUENCE FROM N.A.				
RC	STRAIN=RN4282;				
RA	Barry P.C., Novick R.P.;				
RL	Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; U93688; AAL67620.1; -				
DR	Interpro; IPR001961; Staph_Strep_toxin.				
DR	Pfam; PF011123; Staph_Strp_toxin; 1.				
DR	Dfam; DF02876; Staph_Strp_tox C; 1.				
DR	PRINTS; PR00279; BACTRLTOXIN				
DR	PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.				
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_2; UNKNOWN_1.				
SQ	SEQUENCE 256 AA; 29794 MW; 9E2FI3790823A/DF CRC64;				
Query Match 22.9%; Score 278.5; DB 2; Length 256;					
Best Local Similarity 29.3%; Pred. No. 2.3e-15;					
Matches 68; Conservative 34; Mismatches 69; Indels 61; Gaps					
QY	17 ELQGTALGNLK--QIYYYSKAITSEKSADQLNTLLFKGFTTGHHPWNOLLVDLGS 73	:	:	:	:
Db	58 KLQCVSSGNFSTHQLEYIDGKYTLVSYQ-----FNH-----	:	:	:	:
QY	74 TAATSEYEGS-----SDVLGVAYGYCAGTGPKNKTACMYGGVTLDHNNRLTEKKVPIN 128	:	:	:	:
Db	89 -----EYEAKELDKHKVDIFGISYGLC-----NTKYMGGITLANQN-LDKPRNIPIN 136	:	:	:	:
QY	129 LWIDKGQTTPIDPKVTSKRKEVTVQELDQAARHLHGKFGLYN-----SDSFG 176	:	:	:	:
Db	137 LWANGKNTISTOKVSQTKREVTAQEIDILRKYLQNEYNIYGFNKTKKGQEYGYSKEN 196	:	:	:	:
QY	177 GKVGROGLIVPHSSEGSVSYDLDFAQQOYPDTLRLRYRNTTISSTLSISL 228	:	:	:	:
Db	197 SGFNKNGKITFLHNNEPSTFYDLFTGTGQAESFLKIYNNDKTIDAENFHLDV 248	:	:	:	:
RESULT 22					
ID	Q06532	PRELIMINARY;	PRT;	239 AA.	
AC	Q06532;				
DT	01-NOV-1996	(TrEMBLrel. 01, Created)			
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)			
DT	01-MAR-2002	(TrEMBLrel. 20, Last annotation update)			
DE	Enterotoxin type C-7 (SEC740N) (Fragment).				
OS	Staphylococcus aureus.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;				
OC	Staphylococcus.				


```
Matches 74; Conservative 48; Mismatches 102; Indels 19; Gaps 9;
QY 2 EKSEELNEXDLRKSLQGTALGNLKOIYYVNSKAITSSSE-KSAQDFLTNTLLFKGFFTG 60
Db 1 ESQPDTPDELHKASKFTG-LWENKVL--YDRYVSATKVSVDKFLAHLDIYINLSDKK 57
QY 61 HPWYNLLVLDGTAATSEYSGSSVDLYGAYGYQC-----AGTTPNKACTMYGGVTL 113
Db 58 LKNYDKVKTELLNEDLAKKYKDEVDVGSYVYVNCYFSSKDNVGVKVTGGKTCMYGGITK 117
QY 114 HNNRLTEK--KVPINLWIDGKQTTPIDKVKTSKEYTVQELDLQARHYLHGFLYN 171
Db 118 HEGNFDNGKLNQVLRVY-ENKRNITSE-VQTDKKSQAQELDIKARNFLINKNLVE 175
QY 172 SDSFGKVGORGLIVFHSSGVSVDLFDQAQGYPD--TLLRIYRDNNTTISSTLSISLY 229
Db 176 FNS--SPYGTGYIKFIENNGNTFWIDMPAPGDKFDQSKYLMYNDKNTVDSKSKVIEVH 233
QY 230 LYT 232
Db 234 LTT 236

RESULT 28
O54476
ID O54476 PRELIMINARY; PRT; 242 AA.
AC O54476;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-WAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enterotoxin.
GN ENT.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=RN4282; TRANSPOSON=TN557;
RX MEDLINE=98385824; PubMed=9720870;
RA Lindsay J.A., Ruzin A., Ross H.F., Kurepina N., Novick R.P.;
RT The gene for toxic shock toxin is carried by a family of mobile
RT pathogenicity islands in Staphylococcus aureus.;
RL Mol. Microbiol. 29:527-543(1998).
DR EMBL: U93688; AAC28968.1; -.
DR HSP: P13163; 1SXT.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
SQ SEQUENCE 242 AA: 27720 MW; 28CB4F2EB6B66B9 CRC64;

Query Match 21.6%; Score 263; DB 2; Length 242;
Best Local Similarity 30.1%; Pred. No. 4.le-14;
Matches 66; Conservative 35; Mismatches 92; Indels 26; Gaps 7;
QY 23 LGNLKOIYYVNSKAITSSSEKSAQDFLTNTLLFKGFFTHPWYNLLVLDGTAATSEYEG 82
Db 29 IDNLRNFYTKKDFVDLKDVKDNDTPIANOLQF-----SNESY-DLISESKDFNKSFKG 82
QY 83 SSVDLYGAYGYQCAGGTPNKACTMYGGVTLHDNNRLTEKKVPIINLWIDGKQTTPIDK 142
Db 83 KKLDFEGISYNGQC-----NTKYIYGVTA-TNEYLDSRNIPINIWINGNHKTISTNK 135
QY 143 VKTSKEVTQVELDLQARHYLHGKFGLY-----NSDSFGK-----VORGLIVFHSSE 190
Db 136 VSTNKKLVTAQEIIDLKRLKYLQEEYNIYHNGTKGEEYGHKSKFYSGFNIGKVTFLHNN 195
QY 191 GSTVSYDLF-DAQGVYPTDLLRIYRDNNTTISSTLSISL 228
Db 196 NDTFSYDLFTGDDGLPKSFLKIYEDNKTVESEKPHLDV 234

RESULT 30
Q06534
ID Q06534 PRELIMINARY; PRT; 239 AA.
AC Q06534;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enterotoxin type C (SECOVINE) (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=94011313; PubMed=8406814;
RA Marr J.C., Lyon J.D., Roberson J.R., Lupton M., Bohach G.A.;
RT Characterization of novel type C staphylococcal enterotoxins:
RT biological and evolutionary implications.;
RL Infect. Immun. 61:4254-4262(1993).
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- SIMILARITY: THE DIFFERENT S. AUREUS ENTEROTOXINS AND S. PYOGENES
CC EMBL; L13379; AAA26623.1; -.
DR HSP: P34071; 1SE2.
```

```
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_strp_toxin; 1.
DR Pfam; PF02876; Staph_strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Superantigen.
FT NON_TER 1
SQ SEQUENCE 239 AA; 27517 MW; F354742619C8D196 CRC64;

Query Match 21.2%; Score 258.5; DB 2; Length 239;
Best Local Similarity 30.0%; Pred. No. 9.5e-14;
Matches 73; Conservative 48; Mismatches 103; Indels 19; Gaps 9;

QY 2 EKSEETNEKDLRKKSLQGTALGNLKOIYYYNKSKAITSSSE-KSADQFLNTLLFKGFFTG 60
Db 1 ESQPDPTDELHASKFTG-LNENKVL--YDRVYSATKVKSVKDFLAHDLIYNISDKK 57

QY 61 HPWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQC-----AGTTPNKTKACTMYGGVTL 113
Db 58 LKNYDKVKTTELLNEDLAKKDEVDVYGSNYVNCFSKDNVGVKVTGKTCMYGGITK 117

QY 114 HDNNRLTEE--KKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYN 171
Db 118 HEGNHFQNGNLQNLVRY-ENKRNITISFE-VQTDKKSVAQELDIKARSFLINKNLYE 175

QY 172 SDSFGKGVORGLIVFHSSEGSVSYDLFAQGOYPD--TLLRIYRDNNTTISSTLSISLY 229
Db 176 FNS--SPYETGYIKFTENNGNTFWDMPPAGDKFDOSKYLMMYNDKNTVDSKSVKIEVH 233

QY 230 LYT 232
Db 234 LTT 236

RESULT 31
P97163 PRELIMINARY; PRT; 236 AA.
AC P97163; P97164;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Type A exotoxin precursor (Fragment).
GN SPEA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-MGAS156, AND MGAS500;
RX MEDLINE=92044323; PubMed=1940804;
RA Nelson K., Schlievert P.M., Selander R.K., Musser J.M.;
RT "Characterization and clonal distribution of four alleles of the speA
RT gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
RT Streptococcus pyogenes.";
RL J. Exp. Med. 174:1271-1274(1991).
DR EMBL; X61556; CAA43754.1; -.
DR EMBL; X61557; CAA43755.1; -.
DR EMBL; X61558; CAA43756.1; -.
DR EMBL; X61559; CAA43757.1; -.
DR EMBL; X61554; CAA43752.1; -.
DR HSP; P08095; IB12.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_strp_toxin; 1.
DR Pfam; PF02876; Staph_strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal.
FT NON_TER 1 1
```

```
FT SIGNAL <1 22 POTENTIAL.
FT CHAIN 23 >236 TYPE A EXOTOXIN.
FT NON_TER 236
SQ SEQUENCE 236 AA; 27454 MW; 3FB3F41ABDC1A84 CRC64;

Query Match 20.9%; Score 254.5; DB 2; Length 236;
Best Local Similarity 32.0%; Pred. No. 2e-13;
Matches 74; Conservative 44; Mismatches 94; Indels 19; Gaps 10;

QY 4 SEINEKDLRKKSELOGTAL-GNLKOIYY-YNSKAIT--SSEKSADQFLNTLLFKGFFTG 60
Db 17 SQEVEFAQQDPDSQLHRSLSLVKNLQNIYFLYEGDPVTHENVKSVQDQLLSHDLIYN--VS 73

QY 61 HPWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQC--AGTTPNKTKACTMYGGVTLHDNNRL 119
Db 74 GPNYDKLKTENLQENMATLFKDNVDIYGVYHLCYLCEAERSACIYGGVYNHSGNHL 133

QY 120 TEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYNSDSFGKV 179
Db 134 EIPKKIVKVSIDGIO-SLSFD-IETNKMVTAQELDYKRVKYLTDNKQLYTNGP--SKY 189

QY 180 ORGLIVFHSSEGSVSYDLFD---AOGQYPTDLLRIYRDNNTTISSTLSI 226
Db 190 ETGYIKFIPKNKESEWFDFPEPEFTQSKY----LMIYKDNETLDSNTSQI 236

RESULT 32
Q54779 PRELIMINARY; PRT; 236 AA.
AC Q54779; Q54613; Q54736; Q54740; Q54741;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Type A exotoxin precursor (Fragment).
GN SPEA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-MGAS624 AND MGAS158 AND MGAS485 AND MGAS491, AND MGAS495;
RX MEDLINE=92044323; PubMed=1940804;
RA Nelson K., Schlievert P.M., Selander R.K., Musser J.M.;
RT "Characterization and clonal distribution of four alleles of the speA
RT gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
RT Streptococcus pyogenes.";
RL J. Exp. Med. 174:1271-1274(1991).
DR EMBL; X61569; CAA43767.1; -.
DR EMBL; X61572; CAA43770.1; -.
DR EMBL; X61568; CAA43766.1; -.
DR EMBL; X61570; CAA43768.1; -.
DR EMBL; X61571; CAA43769.1; -.
DR HSP; P08095; IB12.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_strp_toxin; 1.
DR Pfam; PF02876; Staph_strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 22 POTENTIAL.
FT CHAIN 23 >236 TYPE A EXOTOXIN.
FT NON_TER 236
SQ SEQUENCE 236 AA; 27468 MW; 29DF2AD575623A84 CRC64;

Query Match 20.8%; Score 253.5; DB 2; Length 236;
Best Local Similarity 31.6%; Pred. No. 2.4e-13;
Matches 73; Conservative 45; Mismatches 94; Indels 19; Gaps 10;

QY 4 SEINEKDLRKKSELOGTAL-GNLKOIYY-YNSKAIT--SSEKSADQFLNTLLFKGFFTG 60
Db 17 SQEVEFAQQDPDSQLHRSLSLVKNLQNIYFLYEGDPVTHENVKSVQDQLLSHDLIYN--VS 73
```

```
Db 17 SOEFAQQDPDPSQLHRSSLVKNLQNIYELGDPVTHENVKSVQDQLSHDLIYN--VS 73
Qy 61 HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYYGQC-AGGTPNKTKACMYGGVTLHDNNRL 119
Db 74 GPNYDKLKTTELKNOEMATLFFKDNKDIIYGVYHYHLYCILENERSACIYGGVTNHEGNHL 133
Qy 120 TEEKKVPINLWDGKQTTVPIDKVKTSKEVTQVQELDQARHYLHGKFGLYNSDSFGGKV 179
Db 134 EIPKKIVKVSIDGIQ-SLSFD-IETNKKMVTQAQLDYKVRKYLTDNKLQLYTNGP--SKY 189
Qy 180 QRGLIVFHSSEGSVSYDLFD----AQGYPTLLRIYRDNNTTISSTLSI 226
Db 190 ETGYIKFIPKNKESEWFDFPEFTQSKY----LMIYKDNETLDSNTSQI 236

RESULT 33
Qy 09R931 PRELIMINARY; PRT; 222 AA.
AC 09R931;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Exotoxin A (Fragment).
GN SPEA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D709;
RX MEDLINE=99137798; PubMed=9952369;
RA Bessen D.E., Izso M.W., Fiorentino T.R., Caringal R.M.,
RA Hollingshead S.K., Beall B.;
RT "Genetic linkage of exotoxin alleles and emm gene markers for tissue
tropism in group A streptococci."
RL J. Infect. Dis. 179:627-636(1999).
DR EMBL; AF055698; AAD1624.1; -.
DR HSSP; P08095; 1B12.
DR InterPro; IPR001961; Strep_toxin.
DR Pfam; PF01123; Strep_toxin; 1.
DR Pfam; PF02876; Strep_toxin_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT NON_TER 1
FT NON_TER 222
FT SEQUENCE 222 AA; 25759 MW; 48BB7ADDCD91FBA3 CRC64;

Query Match 20.7%; Score 252.5; DB 2; Length 222;
Best Local Similarity 32.0%; Pred. No. 2.7e-13;
Matches 73; Conservative 44; Mismatches 92; Indels 19; Gaps 10;

Qy 4 SEINEKDLRKSELOGTAL-GNLKQIY-YNSKAIT-SSEKSADQFLTNTLLKGFPTG 60
Db 5 SOEFAQQDPDPSQLHRSSLVKNLQNIYELGDPVTHENVKSVQDQLSHDLIYN--VS 61
Qy 61 HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYYGQC-AGGTPNKTKACMYGGVTLHDNNRL 119
Db 62 GPNYDKLKTTELKNOEMATLFFKDNKDIIYGVYHYHLYCILENERSACIYGGVTNHEGNHL 121
Qy 120 TEEKKVPINLWDGKQTTVPIDKVKTSKEVTQVQELDQARHYLHGKFGLYNSDSFGGKV 179
Db 122 EIPKKIVKVSIDGIQ-SLSFD-IETNKKMVTQAQLDYKVRKYLTDNKLQLYTNGP--SKY 177
Qy 180 QRGLIVFHSSEGSVSYDLFD----AQGYPTLLRIYRDNNTTISSTS 223
Db 178 ETGYIKFIPKNKESEWFDFPEFTQSKY----LMIYKDNETLDSNT 221

RESULT 34
Qy 054696 PRELIMINARY; PRT; 236 AA.
ID 054696
```

```
AC Q54696;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Type A exotoxin precursor (Fragment).
GN SPEA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS156;
RX MEDLINE=92044323; PubMed=1940804;
RA Nelson K., Schlievert P.M., Selander R.K., Musser J.M.;
RT "Characterization and clonal distribution of four alleles of the speA
gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
Streptococcus pyogenes."
RL J. Exp. Med. 174:1271-1274(1991).
DR EMBL; X61573; CAA43771.1; -.
DR HSSP; P08095; 1B12.
DR InterPro; IPR001961; Strep_toxin.
DR Pfam; PF01123; Strep_toxin; 1.
DR Pfam; PF02876; Strep_toxin_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal.
FT NON_TER 1
FT SIGNAL <1 22 POTENTIAL.
FT CHAIN 23 >236 TYPE A EXOTOXIN.
FT NON_TER 236 236
FT SEQUENCE 236 AA; 27575 MW; 70F54120E79127DF CRC64;

Query Match 20.6%; Score 251.5; DB 2; Length 236;
Best Local Similarity 31.7%; Pred. No. 3.6e-13;
Matches 72; Conservative 42; Mismatches 102; Indels 11; Gaps 7;

Qy 4 SEINEKDLRKSELOGTAL-GNLKQIY-YNSKAIT-SSEKSADQFLTNTLLKGFPTG 60
Db 17 SOEFAQQDPDPSQLHRSSLVKNLQNIYELGDPVTHENVKSVQDQLSHDLIYN--VS 73
Qy 61 HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYYGQC-AGGTPNKTKACMYGGVTLHDNNRL 119
Db 74 GPNYDKLKTTELKNOEMATLFFKDNKDIIYGVYHYHLYCILENERSACIYGGVTNHEGNHL 133
Qy 120 TEEKKVPINLWDGKQTTVPIDKVKTSKEVTQVQELDQARHYLHGKFGLYNSDSFGGKV 179
Db 134 EIPKKIVKVSIDGIQ-SLSFD-IETSKMVTQAQLDYKVRKHLTDNKLQLYTNGP--SKY 189
Qy 180 QRGLIVFHSSEGSVSYDLFDQAQGYPTLLRIYRDNNTTISSTLSI 226
Db 190 ETGYIKFISKDETWFDFPEFPEFNOVKYLMYKDNETLDSSTSQI 236

RESULT 35
Qy 054739 PRELIMINARY; PRT; 260 AA.
AC 054739; Q54737;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE SUPERANTIGEN SSA.
GN SSA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94222556; PubMed=8168951;
RA Reda K.B., Kapur V., Mollick J.A., Lamphear J.G., Musser J.M.,
RA Rich R.R.;
```

SQ SEQUENCE 260 AA; 29797 MW; 2DD96017DE9D4F49 CRC64;

Query Match 20.6%; Score 251; DB 2; Length 260;
Best Local Similarity 31.2%; Pred. No. 4.5e-13;
Matches 72; Conservative 44; Mismatches 97; Indels 18; Gaps 9;

OY 10 KDLRKKSELOCTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTGHWPWNLLV 69
: : | : | : ||| : | : : : : | : | : :
Db 35 EQLNKSSQFTG-VMGNLRL-LDNHFEVTNVRSQTGLQLHDLPFIKDLKLKNYDSVKT 92

OY 70 DLGSTAATSEYGSSVDLYGAYYQCAGGTPN-----KTACMYGGVTLHDNNRLTEKK 124
: : | : | : ||| : | : : : : | : | : :
Db 93 EFNSKDLATYKNDVPDFGSNYCYSGNSCKNAKTCMTGGVTEHRNQI--EKG 150

OY 125 VPINLWI---DGKOTTVPIDRVKTSKRETVQELDQARHYLHGKFGLYNSDFCGKVQR 181
: : | : | : : : : : | : | : | : | : | : :
Db 151 FP-NITVKVVEDNEIISFD-ITTNKKQVTVQVELDKTRKILVSRKNLIYEFNN--SPIET 206

OY 182 GLIVFHSSSEGSTVSDFLDAGQQYPD--TLRIYRDNTTISTSLISLYL 230
: : | : | : ||| : | : : : : | : | : : : :
Db 207 GYIKFIESSGDSFWYDMPPAPGAIFDQSGLMLYNDKNKTVSSSAIAIEVHL 257

RESULT 37

Q54738 PRELIMINARY; PRT; 260 AA.

ID Q54738 AC Q54738

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Superantigen SSA.

GN SSA.

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;

OC Streptococcaceae; Streptococcus.

ON NCBI_TaxID=1314;

OX [1]

PC SEQUENCE FROM N.A.

RP STRAIN=MGAS 1842;

RX MEDLINE=94222556; PubMed=8168951;

RA Reda K.B., Kapur V., Mollick J.A., Lamphear J.G., Musser J.M., Rich R.R.;

RA Reda K.R.;

RT "Molecular characterization and phylogenetic distribution of the streptococcal superantigen gene (ssa) from Streptococcus pyogenes.";

RL Infect. Immun. 62:1867-1874(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MGAS 1842;

RX MEDLINE=96178602; PubMed=8606073;

RA Reda K.B., Kapur V., Goela D., Lamphear J.G., Musser J.M., Rich R.R.;

RT "Phylogenetic distribution of streptococcal superantigen SSA allelic variants provides evidence for horizontal transfer of ssa within Streptococcus pyogenes.";

RL Infect. Immun. 64:1161-1165(1996).

DR EMBL; U48793; AAB02149.1; -

DR HSP; P01552; ISEB.

DR InterPro; IPR001961; Staph/Strep_toxin.

DR Pfam; PF011123; Staph_Strep_toxin; 1.

DR Pfam; PF02876; Staph_Strep_tox_C; 1.

DR PRINTS; PS00279; BACTRLTOXIN.

DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.

DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.

SQ SEQUENCE 260 AA; 29836 MW; C122141693B42AD6 CRC64;

Query Match 20.6%; Score 251; DB 2; Length 260;
Best Local Similarity 31.2%; Pred. No. 4.5e-13;
Matches 72; Conservative 44; Mismatches 97; Indels 18; Gaps 9;

OY 10 KDLRKKSELOCTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTGHWPWNLLV 69
: : | : | : ||| : | : : : : | : | : :
Db 35 EQLNKSSQFTG-VMGNLRL-LDNHFEVTNVRSQTGLQLHDLPFIKDLKLKNYDSVKT 92

OY 70 DLGSTAATSEYGSSVDLYGAYYQCAGGTPN-----KTACMYGGVTLHDNNRLTEKK 124

DT	01-WAR-2002 (TReMBLrel. 20, Last annotation update)	
DE	Pyrogenic exotoxin A (Fragment).	
GN	SPeA.	
OC	Streptococcus equisimilis.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;	
OC	Streptococcaceae; Streptococcus.	
OC	NCBI_TaxID=119602;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=4951;	
RT	Kalia A., Bessen D.E.;	
RA	"Presence of streptococcal pyrogenic exotoxin A and C genes in human	
RT	isolates of group G Streptococci."	
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; A1049745; AAL06068.1; -	
DR	InterPro; IPR001961; Strep/Strep-toxin.	
DR	Pfam; PF011123; Strep_strep_toxin; 1.	
DR	Pfam; PF02876; Strep_strep_toxin; 1.	
DR	PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.	
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_2; UNKNOWN_1.	
FT	NON_TER	1
FT	NON_TER	222
FT	SEQUENCE	222 AA; 25884 MW; 121F8460992818F8 CRC64;
QY	Query Match	30.5%; Score 250; DB 2; Length 222;
QY	Best Local Similarity	32.9%; Pred. No. 4.4e-13;
QY	Matches	74; Conservative 40; Mismatches 99; Indels 12; Gaps
QY	4	SEINEKDLRKKELOGTAL-GNLKQIY--YYNSKAITSEKSAQDLFWTLFLPKGFFTG 60
QY	5	SQEVFAQQDPNPQLHRSSLLKQLNIYFLYEGDPVVHENKVSQQLLSDLIYN--VS 61
QY	61	HPWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQC-AGTGNKTKACMGYGVTLHDNNRL 119
QY	62	GLNYDKLTKLNKREMTLEFKNNVDIYGYEYYHCYLCRNARRACIYGVTHHEGNHL 121
QY	120	TEKKVPINLWDGKQTPIDPKVYKTSKKEVTVQELDLQARHYLHGKFGYNSDSFGKV 179
QY	122	EIPKNILVKVSDIGQ-SLSFD-IETSKMVTAEQLDYKVRKHLTDNNQLYNGP--SKY 177
QY	180	QRGLIVPHSSEGSTVSVLDLDAOQYDPDTLLRIYRONTTI-SSTS 223
QY	178	ETGVKIFISKDKETFFDFPEPFQNVQKYLMIYKDNETLDSSTS 222
RESULT	40	
Q92NF2	PRELIMINARY;	PRT; 258 AA.
ID	Q92NF2	
AC	Q92NF2;	
DT	01-MAY-1999 (TReMBLrel. 10, Created)	
DT	01-MAY-1999 (TReMBLrel. 10, Last sequence update)	
DT	01-MAR-2002 (TReMBLrel. 20, Last annotation update)	
DE	Enterotoxin type Gv.	
GN	SEG2.	
OS	Staphylococcus aureus.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;	
OC	Staphylococcus.	
OC	NCBI_TaxID=1280;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	MEDLINE=20260630; PubMed=10803494;	
RA	Abe J., Ito Y., Onimaru M., Kohsaka T., Takeda T.;	
RT	"Characterization and distribution of a new enterotoxin-related	
RL	superantigen produced by Staphylococcus aureus."	
RL	Microbiol. Immunol. 44:79-88 (2000).	
DR	EMBL; AB016487; BAA36693.1; -	
DR	HSSP; P01552; 1SBB.	
DR	InterPro; IPR001961; Strep/Strep-toxin.	
DR	Pfam; PF011123; Strep_strep_toxin; 1.	
DR	Pfam; PF02876; Strep_strep_toxin; 1.	
DR	PRINTS; PR00279; BACTRLTOXIN.	
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_1.	
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.	
SQ	SEQUENCE	258 AA; 29839 MW; 1229246D83F4FB77 CRC64;

Db 17 SOEFAQDQDP SOLHRSSLVKQLNIYFLYEGDPVTHENVKSDQLLSDLIYN---VS 73
 QY 61 HPWYNLLVDLVDGTAATSEYEGSSVDLYGAYGYQC--AGGTPNKATCMYGGVTLHDNNRL 119
 Db 74 GPNYDKLKTTELKQEMATLEFKDKRNDVIYSVEYHLCYLCENASRSACIYGGVTNHGNNHL 133
 QY 120 TEEKKVPINLWIDGKQTTVPIDPKVKTSKKEVTVOELDQARHYLHGKFGLYNSDSFGGKV 179
 Db 134 EIPKKIVVKVSDIGQ--SLSFQ--IETNKKMTAQELDYKVRKVLTONKQLYNGP--SKY 189
 QY 180 ORGLIVPHSSEGSTVSVDLFD---AQGOYPTDLLRIYRDNTTISSTLSI 226
 Db 190 ETGYIKETIPKNKESFWDFPEPEFTQSKY----LMIYKDNELTDSNTSOI 236

RESULT 42
 ID Q8RR77 PRELIMINARY; PRT; 233 AA.
 AC Q8RR77;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Enterotoxin G (Fragment).
 GN SEG.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OC NCBI_TaxID=1280;
 RP [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=21871379; PubMed=11880405;
 RA Omoe K., Ishikawa M., Shimoda Y., Hu D.L., Ueda S., Shinagawa K.;
 RT "Detection of seg, seh, and sei genes in Staphylococcus aureus
 RT isolates and Determination of the Enterotoxin Productivities of S.
 RT aureus Isolates Harboring seg, seh, or sei Genes.";
 RL J. Clin. Microbiol. 40:857-862(2002).
 DR EMBL; AB060535; BAB85989.1; -;
 FT NON_TER 1
 SQ SEQUENCE 233 AA; 27040 MW; BCE85287D63BF60D CRC64;

Query Match 19.8%; Score 241; DB 2; Length 233;
 Best Local Similarity 28.8%; Pred. No. 2 6e-12;
 Matches 72; Conservative 48; Mismatches 88; Indels 42; Gaps

Qy 3 KSEINEKDLRKSELOCTALGNLKKIYY---YNSKAITSEKSAQOFLTNTLLFKGFFT 59
 Db 5 KLDELNKVSDYKNN--KGT-MGNVMNLYTSPPEVGRGVINSR---QFLSHDLIFP---I 54
 Qy 60 GHPWYNLLVDLGDSTATSEYEGSSVDLYGAYGYQCA-----GGTPNKKTACH 107
 Db 55 EYKSYNEVKTELENTELANNYKDKVDIFGVFPVYTCIIPKSEPDINQFPG-----CCM 109
 Qy 108 YGGVTLH--DNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDQARHYLHG 165
 Db 110 YGGLTFNSENER--DKLITVQVTDNRSLG--FTITTKNNVYIQELDYKARHLTK 164
 Qy 166 KFGLYNSDSFGGKVQRGLIVPHSSEGSTVSVDLFDAGQGYPT---LLRIYRDNTTISST 222
 Db 165 EKKLYEFD--GSAPESGYIKFTTEKNNTSFWFOLFPPKELVPFPYKFLNIYGDKNVVDK 222

Qy 223 SLSISLYLT 232
 Db 223 SIRKEVELNT 232

RESULT 43
 ID Q9EZM3 PRELIMINARY; PRT; 258 AA.
 AC Q9EZM3;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE SEGL29P.
GN SEGL29P.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A900322;
RX MEDLINE=20571956; PubMed=11123352;
RA Jarraud S., Peyrat M.A., Lim A., Tristan A., Bes M., Mougel C.,
RA Etienne J., Vandenesch F., Bonneville M., Lina G.;
RT "esc, A highly prevalent operon of enterotoxin gene, forms a putative
RT nursery of superantigens in Staphylococcus aureus.";
RL J. Immunol. 166:569-577(2001).
DR EMBL; AF285760; AAG36957.1; -.
DR HSSP; P01552; 1SBB.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_tox_C; 1.
DR PRINTS; PR00279; BACTRUTOXIN.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
SQ SEQUENCE 258 AA; 29956 MW; A85619E2FE21E3B7 CRC64;

Query Match 19.8%; Score 241; DB 2; Length 258;
Best Local Similarity 28.8%; Pred. No. 3e-12;
Matches 72; Conservative 48; Mismatches 88; Indels 42; Gaps 12;

QY 3 KSEETNEKDLRKSELOCTALGNLKOIYY---YNSKAITSEKSADQFLTNLLFKGFET 59
Db 30 KLELNKVS DYKNN--KGT-MGNVNNLYTSPVVEGRGVNSR---QFLSHDLIFP---I 79
QY 60 GHPWYNDLLVLGSGTAATSEYEGSSVDLYGAYGYQCA-----GGTPNKTACM 107
Db 80 EYKSYNEVKTELENTALEANNYKDKKVDIFGVFYFTCIIPKSEPDINQFEG---CCM 134
QY 108 YGVVTLH--DNNRLTEERKVPINLWDGKQTVPIDKVKTSKEVTVOELDLQARHYLHG 165
Db 135 YGLTFNSENER---DKLITVQVITDRQSLG--FTITTKNMVYTIQELDYKARHWLTK 189
QY 166 KFGLYNSDFGKVGORGLIVFHSSEGSVSYVDLFDQAQGYPDT---LLRIYRDNFTTIST 222
Db 190 EKKLYEFD--GSAFESGYIKTEKNNTSFWDLPFKKELVPVPYKFLNIYGDKNVDSK 247
QY 223 SISISLYLYT 232
Db 248 SIKMEVFLNT 257

RESULT 44
ID 085382 PRELIMINARY; PRT; 258 AA.
AC 085382;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Extracellular enterotoxin type G precursor.
GN SEG OR S4V1824 OR S41642.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. aureus; STRAIN=FR1572;
RX MEDLINE=98298056; PubMed=9632603;
RA Munson S.H., Tremaine M.T., Betley M.J., Welch R.A.;
RT "Identification and characterization of staphylococcal enterotoxin
RT types G and I from Staphylococcus aureus.";
RL Infect. Immun. 66:3337-3348(1998).
```

```
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Hiramatsu K.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF064773; AAC26660.1; -.
DR EMBL; AP003363; BAB57986.1; -.
DR EMBL; AP003135; BAB42910.1; -.
DR HSSP; P01552; 1SBB.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_tox_C; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal; Complete proteome.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 258 ENTEROTOXIN TYPE G.
SQ SEQUENCE 258 AA; 29940 MW; E2982101701D012C CRC64;

Query Match 19.8%; Score 241; DB 16; Length 258;
Best Local Similarity 28.8%; Pred. No. 3e-12;
Matches 72; Conservative 48; Mismatches 88; Indels 42; Gaps 12;

QY 3 KSEETNEKDLRKSELOCTALGNLKOIYY---YNSKAITSEKSADQFLTNLLFKGFET 59
Db 30 KLELNKVS DYKNN--KGT-MGNVNNLYTSPVVEGRGVNSR---QFLSHDLIFP---I 79
QY 60 GHPWYNDLLVLGSGTAATSEYEGSSVDLYGAYGYQCA-----GGTPNKTACM 107
Db 80 EYKSYNEVKTELENTALEANNYKDKKVDIFGVFYFTCIIPKSEPDINQFEG---CCM 134
QY 108 YGVVTLH--DNNRLTEERKVPINLWDGKQTVPIDKVKTSKEVTVOELDLQARHYLHG 165
Db 135 YGLTFNSENER---DKLITVQVITDRQSLG--FTITTKNMVYTIQELDYKARHWLTK 189
QY 166 KFGLYNSDFGKVGORGLIVFHSSEGSVSYVDLFDQAQGYPDT---LLRIYRDNFTTIST 222
Db 190 EKKLYEFD--GSAFESGYIKTEKNNTSFWDLPFKKELVPVPYKFLNIYGDKNVDSK 247
QY 223 SISISLYLYT 232
Db 248 SIKMEVFLNT 257

RESULT 45
Q936G4
ID Q936G4 PRELIMINARY; PRT; 259 AA.
AC Q936G4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Enterotoxin-like protein.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M;
RX MEDLINE=95050273; PubMed=7961465;
RA Lin W.S., Cunneen T., Lee C.Y.;
RT "Sequence analysis and molecular characterization of genes required
RT for the biosynthesis of type I capsular polysaccharide in
RT Staphylococcus aureus.";
RL J. Bacteriol. 176:7005-7016(1994).
```



```
RESULT 48
Q9LAD8
ID Q9LAD8 PRELIMINARY; PRT: 209 AA.
AC Q9LAD8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-WAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mitogenic exotoxin z-7 (Fragment).
GN SMEZ-7.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=11574;
RX MEDLINE=20273982; PubMed=10811869;
RA Prof T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
RA Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
RT Mosaic Structure, and Significant Antigenic Variation.";
RL J. Exp. Med. 191:1765-1776(2000).
DR EMBL; AF143657; AAF66657.1; -.
DR HSP; P13163; ISXT.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_strep_toxin; 1.
DR Pfam; PF02876; Strep_strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; BACTRLTOXIN.
DR NON_TER 1
FT SEQUENCE 209 AA; 24064 MW; 67E977CDSAA934F7 CRC64;
SQ SEQUENCE 209 AA; 24064 MW; 67E977CDSAA934F7 CRC64;

Query Match 18.0%; Score 219; DB 2; Length 209;
Best Local Similarity 29.1%; Pred. No. 1.6e-10;
Matches 58; Conservative 39; Mismatches 66; Indels 36; Gaps 6;

QY 50 NTLFLKGFPTGHPWYNDLLVD-----LGSTAATSEYEGSSVDLYGAY----- 91
Db 6 NSLLRNIYSTIVYVSDTLIDFKTSHNLVTKKLDVRDARDFINSEMDEYAANDFKAGDK 65
QY 92 -----YGQCAGGTPNKTYGVTLDHNNRLTEKKVPINLWIDGKQTTVPIDK 142
Db 66 IAVFSVPFDWNYLSEG---KVATYTYGGITPYQKTSI--PKNIPVNLWNGKQISVPYNE 120
QY 143 VKTSKEVTVOELDLQARHLYHGKFGLYNSDSFGGKVORGLIVFHSSEGS-TVSYDLFDA 201
Db 121 ISTNKTVTVAQEIDLKVRKFLIAQHOLYSS---GSSYKSGKLIVFHTNDSKYSFDLFYT 177
QY 202 OGQYDPDLLRIYRDNNTIS 220
Db 178 GYRDKESIFKVKDNKSNF 196

RESULT 49
Q9LAD1
ID Q9LAD1 PRELIMINARY; PRT: 209 AA.
AC Q9LAD1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mitogenic exotoxin z-14 (Fragment).
GN SMEZ-14.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4202;
RX MEDLINE=20273982; PubMed=10811869;
RA Prof T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
RA Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
RT Mosaic Structure, and Significant Antigenic Variation.";
RL J. Exp. Med. 191:1765-1776(2000).
DR EMBL; AF143670; AAF66669.1; -.
DR HSP; P13163; ISXT.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_strep_toxin; 1.
DR Pfam; PF02876; Strep_strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; BACTRLTOXIN.
DR NON_TER 1
FT SEQUENCE 209 AA; 24100 MW; 53049A11599BEA68 CRC64;
SQ SEQUENCE 209 AA; 24100 MW; 53049A11599BEA68 CRC64;

Query Match 17.8%; Score 217; DB 2; Length 209;
Best Local Similarity 29.1%; Pred. No. 2.3e-10;
Matches 58; Conservative 38; Mismatches 67; Indels 36; Gaps 6;

QY 50 NTLFLKGFPTGHPWYNDLLVD-----LGSTAATSEYEGSSVDLYGAY----- 91
Db 6 NSLLRNIYSTIVYVSDTLIDFKTSHNLVTKKLDVRDARDFINSEMDEYAANDFKAGDK 65
```

```
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
RT Mosaic Structure, and Significant Antigenic Variation.";
RL J. Exp. Med. 191:1765-1776(2000).
DR EMBL; AF143664; AAF66664.1; -.
DR HSP; P13163; ISXT.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_strep_toxin; 1.
DR Pfam; PF02876; Strep_strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; BACTRLTOXIN.
DR NON_TER 1
FT SEQUENCE 209 AA; 24122 MW; 0CF5D429E1B96EFE CRC64;
SQ SEQUENCE 209 AA; 24122 MW; 0CF5D429E1B96EFE CRC64;

Query Match 17.8%; Score 217; DB 2; Length 209;
Best Local Similarity 29.1%; Pred. No. 2.3e-10;
Matches 58; Conservative 38; Mismatches 67; Indels 36; Gaps 6;

QY 50 NTLFLKGFPTGHPWYNDLLVD-----LGSTAATSEYEGSSVDLYGAY----- 91
Db 6 NSLLRNIYSTIVYVSDTLIDFKTSHNLVTKKLDVRDARDFINSEMDEYAANDFKAGDK 65
QY 92 -----YGQCAGGTPNKTYGVTLDHNNRLTEKKVPINLWIDGKQTTVPIDK 142
Db 66 IAVFSVPFDWNYLSEG---KVATYTYGGITPYQKTSI--PKNIPVNLWNGKQISVPYNE 120
QY 143 VKTSKEVTVOELDLQARHLYHGKFGLYNSDSFGGKVORGLIVFHSSEGS-TVSYDLFDA 201
Db 121 ISTNKTVTVAQEIDLKVRKFLIAQHOLYSS---GSSYKSGKLIVFHTNDSKYSFDLFYT 177
QY 202 OGQYDPDLLRIYRDNNTIS 220
Db 178 GYRDKESIFKVKDNKSNF 196

RESULT 50
Q9LAC6
ID Q9LAC6 PRELIMINARY; PRT: 209 AA.
AC Q9LAC6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mitogenic exotoxin z-20 (Fragment)..
GN SMEZ-20.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=10989;
RX MEDLINE=20273982; PubMed=10811869;
RA Prof T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,
RA Fraser J.D.;
RT "The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,
RT Mosaic Structure, and Significant Antigenic Variation.";
RL J. Exp. Med. 191:1765-1776(2000).
DR EMBL; AF143670; AAF66669.1; -.
DR HSP; P13163; ISXT.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_strep_toxin; 1.
DR Pfam; PF02876; Strep_strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; BACTRLTOXIN.
DR NON_TER 1
FT SEQUENCE 209 AA; 24100 MW; 53049A11599BEA68 CRC64;
SQ SEQUENCE 209 AA; 24100 MW; 53049A11599BEA68 CRC64;

Query Match 17.8%; Score 217; DB 2; Length 209;
Best Local Similarity 29.1%; Pred. No. 2.3e-10;
Matches 58; Conservative 38; Mismatches 67; Indels 36; Gaps 6;

QY 50 NTLFLKGFPTGHPWYNDLLVD-----LGSTAATSEYEGSSVDLYGAY----- 91
Db 6 NSLLRNIYSTIVYVSDTLIDFKTSHNLVTKKLDVRDARDFINSEMDEYAANDFKAGDK 65
```

```
Qy 92 -----YGYCAGGTPNKTACMYGGVTLHDNNRLTEEKVPINLWDGKQTTPIDK 142
      :   |   ||| :   |   ||| :   |   ||| :   |   ||| :   |   ||| :
Db 66 IAVESVPEDWNYSKG---KVTAITYGGITPYKTSI--PKNIPVNWLINGKQIPVPYNQ 120

Qy 143 VKTSKKEVTVEQLDLQARHYLHGKFGLYNSDSFGKVGQRGVLIVFHSSEGS-TVSYDLFDA 201
      :|: || ||||: |: |||: |: |||: |: |||: |: |||:
Db 121 ISTNKTTVAQEIDLKVRKFLIAQHQLYSS--GSSYKSGRLVFHTNDNSDKYSFDLPYT 177

Qy 202 QGYPTDLLRIYRDNTTIS 220
      :   :   :   :||| :
Db 178 GYRDKESIIFYVKDKNFSN 196
```

Search completed: January 8, 2003, 11:58:25
Job time : 19.4475 secs

THIS PAGE BLANK (USP)